CAPACIDADES INVESTIGATIVAS

METODOLOGÍA DE LA INVESTIGACIÓN

AUTOR:
GABRIEL OXSA QUISPE

INVESTIGAR HOY, TRANSFORMAR MAÑANA

CAPACIDADES INVESTIGATIVAS

5.1. Planteamiento del problema

5.1.1. Identificación de un problema a partir de la observación de la realidad

- Autor principal: **Fred N. Kerlinger**
- Libro recomendado: *Investigación del Comportamiento*
- Autor de apoyo: **Mario Bunge**
- Obra: *La Ciencia, su Método y su Filosofía*

5.1.2. Formulación de preguntas de investigación derivadas del problema identificado

- Autor principal: **Fidias G. Arias**
- Libro recomendado: *El Proyecto de Investigación*
- Autor de apoyo: **César Bernal**
- Obra: *Metodología de la Investigación*

5.2. Justificación, objetivos e hipótesis de la investigación

5.2.1. Formulación de objetivos general y específicos

- Autor principal: **Fidias G. Arias**
- Libro recomendado: *El Proyecto de Investigación*
- Autor de apoyo: **César Bernal**
- Obra: *Metodología de la Investigación*

5.2.2. Elaboración de la justificación teórica, práctica y metodológica

- Autor principal: **César Augusto Bernal Torres**
- Libro recomendado: *Metodología de la Investigación*

5.2.3. Redacción de hipótesis de investigación e identificación de variables

- Autor principal: **Fred N. Kerlinger**
- Libro recomendado: *Investigación del Comportamiento*
- Autor de apoyo: **Roberto Hernández Sampieri**
- Obra: *Metodología de la Investigación*

5.3. Metodología

5.3.1. Caracterización de la investigación: enfoque, alcance y tipo (Diseño)

- Autor principal: **Roberto Hernández Sampieri**
- Libro recomendado: *Metodología de la Investigación*

5.3.2. Selección de técnicas e instrumentos para la recolección de datos

- Autor principal: **Fidias G. Arias**
- Libro recomendado: *El Proyecto de Investigación*
- Autor de apoyo: **Roberto Hernández Sampieri**
- Obra: *Metodología de la Investigación*

5.4. Análisis y representación de datos

5.4.1. Organización e interpretación de datos cuantitativos y cualitativos – (Organización, procesamiento e interpretación de datos cuantitativos y cualitativos)

- Autor principal: **Roberto Hernández Sampieri**
- Libro recomendado: *Metodología de la Investigación*
- Autor de apoyo: **John W. Creswell**
- Obra: *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*
- **Autor complementario: Mario F. Triola**
- **Obra:** Estadística

5.4.2. Representación de datos mediante tablas y gráficos

- **Autor principal:** Roberto Hernández Sampieri
- **Libro recomendado:** *Metodología de la investigación*
- **Autor de apoyo:** Mario F. Triola
- **Obra:** *Estadística*

5.4.3. Técnicas estadísticas para el análisis de datos

- **Autor principal:** Roberto Hernández Sampieri
- **Libro recomendado:** Metodología de la Investigación
- **Autor de apoyo:** Mario F. Triola
- **Obra:** Estadística
- **Autor complementario:** Richard I. Levin & David S. Rubin
- **Obra:** Estadística para Administración y Economía

5.5. Resultados y discusión

5.5.1. Interpretación de los resultados en relación con los objetivos e hipótesis – (Interpretación y discusión de los resultados en relación con los objetivos, hipótesis y antecedentes)

- Autor principal: **Roberto Hernández Sampieri**
- Libro recomendado: *Metodología de la Investigación*
- Autor de apoyo: **John W. Creswell**
- Obra: *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*

5.5.2. Elaboración de conclusiones fundamentadas y su relación con las hipótesis

- **Autor principal:** César Augusto Bernal Torres
- **Libro recomendado:** *Metodología de la investigación*
- **Autor de apoyo:** Roberto Hernández Sampieri
- **Obra:** *Metodología de la investigación*

5.1. Planteamiento del problema

5.1.1. Identificación de un problema a partir de la observación de la realidad

- Autor principal: **Fred N. Kerlinger**
- Libro recomendado: *Investigación del Comportamiento*
- Autor de apoyo: **Mario Bunge**
- Obra: *La Ciencia, su Método y su Filosofía*

5.1.2. Formulación de preguntas de investigación derivadas del problema identificado

- Autor principal: **Fidias G. Arias**
- Libro recomendado: *El Proyecto de Investigación*
- Autor de apoyo: **César Bernal**
- Obra: *Metodología de la Investigación*

5.1.1. Identificación de un problema a partir de la observación de la realidad

Para el desarrollo del subtema **Identificación de un problema a partir de la observación de la realidad**, se consideran aquellos contenidos que explican el origen del conocimiento científico, el papel de la observación en el proceso de investigación y la identificación de problemas como punto de partida del método científico. Asimismo, se incluyen los criterios para reconocer un problema de investigación y su importancia dentro del proceso científico, conforme a la estructura de los textos de referencia.

N.°	Libros	Capítulo	Página	Contenido
1	Kerlinger, F. N. & Lee, H. B. *Investigación del comportamiento* (4.ª ed.)	Capítulo 1. **La ciencia y el enfoque científico**	p. 13	La investigación científica: definición.
2	Kerlinger, F. N. & Lee, H. B. *Investigación del comportamiento* (4.ª ed.)	Capítulo 1. **La ciencia y el enfoque científico**	p. 14	El enfoque científico.
3	Kerlinger, F. N. & Lee, H. B. *Investigación del comportamiento* (4.ª ed.)	Capítulo 1. **La ciencia y el enfoque científico**	p. 14	Problema – obstáculo – idea.
4	Kerlinger, F. N. & Lee, H. B. *Investigación del comportamiento* (4.ª ed.)	Capítulo 1. **La ciencia y el enfoque científico**	p. 16	Observación – prueba – experimento.
5	Kerlinger, F. N. & Lee, H. B. *Investigación del comportamiento* (4.ª ed.)	Capítulo 2. **Problemas e hipótesis**	p. 21	Problemas.
6	Kerlinger, F. N. & Lee, H. B. *Investigación del comportamiento* (4.ª ed.)	Capítulo 2. **Problemas e hipótesis**	p. 23	Criterios de los problemas y enunciados de problemas.
7	Kerlinger, F. N. & Lee, H. B. *Investigación del comportamiento* (4.ª ed.)	Capítulo 2. **Problemas e hipótesis**	p. 24	Importancia de los problemas e hipótesis.
8	Kerlinger, F. N. & Lee, H. B. *Investigación del comportamiento* (4.ª ed.)	Capítulo 2. **Problemas e hipótesis**	p. 25	Virtudes de los problemas e hipótesis.
9	Mario Bunge. *La ciencia. Su método y su filosofía*	**¿Qué es la ciencia?**	p. 9	Fundamentos de la ciencia como base para comprender la observación de la realidad.
10	Mario Bunge. *La ciencia. Su método y su filosofía*	**¿Cuál es el método de la ciencia?**	p. 49	El método científico como proceso para identificar problemas a partir de la observación de la realidad.

tipo de úlcera fueron tratados con éxito por medio de antibióticos y a un precio considerablemente más bajo. Durante 75 años ningún resultado se tomó como evidencia contra la teoría del ácido-estrés de las úlceras.

La investigación científica: definición

Es más fácil definir la investigación científica que la ciencia. Sin embargo, no sería sencillo lograr un acuerdo de científicos e investigadores en relación con una definición. Aun así intentaremos una: *la investigación científica es una investigación sistemática, controlada, empírica, amoral, pública y crítica de fenómenos naturales. Se guía por la teoría y las hipótesis sobre las presuntas relaciones entre esos fenómenos.* Esta definición requiere poca explicación en tanto que es una declaración condensada y formalizada de muchos aspectos que ya se han presentado o que abordaremos pronto. Sin embargo, es necesario enfatizar dos puntos.

Primero, cuando decimos que la investigación científica es sistemática y controlada queremos decir, de hecho, que la investigación es tan ordenada que los investigadores pueden tener una confianza crítica en los resultados. Como se verá más adelante, las observaciones de la investigación científica son estrictamente disciplinadas. Más aún, entre las muchas explicaciones alternativas de un fenómeno, todas menos una se rechazan de forma sistemática. Así, uno puede tener mayor confianza en que una relación sometida a prueba es tal como es, que si no se hubieran controlado las observaciones y desechado las posibilidades alternativas. En algunos casos es posible establecer una relación de causa-efecto.

Segundo, la investigación científica es empírica. Si el científico cree que algo se da de cierta forma debe demostrarlo de un modo u otro por medio de una prueba independiente externa. En otras palabras, las consideraciones subjetivas deben ser verificadas contra una realidad objetiva. Los científicos siempre deben presentar sus nociones ante el tribunal del cuestionamiento y prueba empíricos. Los científicos son sumamente críticos de sus propios resultados y de los de las investigaciones de los demás. Cada científico que redacta un informe de investigación cuenta con otros científicos que leen lo que él escribe a lo largo de todo el proceso. Aunque es fácil errar, exagerar, generalizar de más al redactar uno mismo su propio trabajo, no es fácil escapar al escrutinio de otros científicos que vigilan la tarea.

En ciencia existe la revisión de pares. Esto significa que otros con igual capacidad y conocimiento son llamados para evaluar el trabajo del científico antes de que sea publicado en revistas especializadas. En este punto hay tantos aspectos positivos como negativos. Es a través de esta revisión de pares que se han descubierto estudios fraudulentos. El ensayo escrito por R.W. Wood (1973) acerca de sus experiencias con el profesor francés Blondlot, sobre la inexistencia de los rayos *N*, aporta una clara demostración de estas revisiones, que resultan buenas para la ciencia y promueven la calidad en la investigación. El sistema, sin embargo, no es perfecto. Hay ocasiones en que la evaluación de pares se ha volcado contra la ciencia. Esto está documentado a través de la historia con personas como Kepler, Galileo, Copérnico, Jenner y Semelweiss. Las ideas de estos individuos no fueron populares entre colegas. Más recientemente, en psicología, el trabajo de John García sobre las restricciones biológicas del aprendizaje fue contrario a los de sus colegas. García consiguió publicar sus hallazgos en una revista (*Bulletin of the Psychonomic Society*) que no tenía evaluación de pares. Algunos investigadores que leyeron y replicaron su trabajo lo encontraron valioso. En la gran mayoría de los casos la revisión de colegas resulta benéfica para la ciencia.

Tercero, el conocimiento obtenido científicamente no está sujeto a una evaluación moral. Los resultados no se juzgan por "malos" ni "buenos", sino en términos de validez y confiabilidad. Sin embargo, el método científico está sujeto a principios de moralidad; es decir, que el científico es responsable de los métodos utilizados para obtener el conocimiento científico. En psicología, los códigos de ética se establecen para proteger a quienes están bajo estudio. La ciencia es una aventura compartida. La formación científica está disponible para todos y el método científico es bien conocido y está a la mano de quienes eligen usarlo.

El enfoque científico

El enfoque científico es una forma especial y sistematizada del pensamiento y del cuestionamiento reflexivos. Dewey (1933/1991), en su influyente trabajo *How We Think (Cómo Pensamos)*, delineó un paradigma general del cuestionamiento. La presente discusión del enfoque científico se basa en gran parte en el análisis de Dewey.

Problema-obstáculo-idea

El científico puede experimentar dificultades para entender, una vaga inquietud acerca de los fenómenos observados y no observados, una curiosidad sobre por qué algo es de la forma en que se presenta. El primer paso —y el más importante— es tener que sacar a la luz una idea, expresar el problema de alguna forma razonablemente manejable. Nunca o rara vez el problema surgirá por completo en esta etapa. El científico debe esforzarse con él, luchar con él y vivir con él. Dewey (1933/1991, p. 108) dice: "Hay una situación problemática, perpleja, irritante, donde la dificultad se encuentra a todo lo largo y ancho de ella, afectándola como un todo." Más tarde o más temprano, explícita o implícitamente, el científico definirá el problema, incluso si su expresión resulta incipiente y tentativa. El científico intelectualiza, como Dewey (p. 109) señala "aquello que al principio es meramente una cualidad *emocional* de toda la situación" (las itálicas son agregadas). En algunos aspectos ésta es la parte más difícil e importante de todo el proceso. Sin algún tipo de definición del problema, el científico difícilmente podrá seguir adelante y esperar que su trabajo sea fructífero. Para algunos investigadores la idea puede provenir de la conversación con un colega, o de la observación de un fenómeno curioso. La idea es que el problema por lo general se inicia con un pensamiento vago o no científico, o con presentimientos no sistemáticos. Después seguirán pasos más refinados.

Hipótesis

Después de intelectualizar el problema, de referirse a experiencias pasadas para posibles soluciones, de observar fenómenos relevantes, el científico puede formular una hipótesis. Una hipótesis es una declaración conjetural, una proposición tentativa acerca de la relación entre dos o más fenómenos o variables. Nuestro científico dirá, "si ocurre tal y tal, entonces resultará tal y tal".

Razonamiento-deducción

Este paso o actividad con frecuencia pasa inadvertido o es poco enfatizado. Quizás es la parte más importante del análisis de Dewey sobre el pensamiento reflexivo. El científico

deduce las consecuencias de la hipótesis que él mismo ha formulado. Conant (1951), al hablar acerca del surgimiento de la ciencia moderna, indica que el nuevo elemento que se aportó en el siglo XVII fue el uso del razonamiento deductivo. Aquí es donde la experiencia, el conocimiento y la perspicacia son importantes.

Con frecuencia el científico, al deducir las consecuencias de una hipótesis formulada, llegará a un problema muy diferente del original. Por otro lado, las deducciones pueden hacer creer que el problema no puede resolverse con las herramientas técnicas actuales. Por ejemplo, antes que la estadística moderna se desarrollara, algunos problemas de investigación del comportamiento eran irresolubles. Era difícil, si no es que imposible, probar dos o tres hipótesis interdependientes de forma simultánea, y también era casi imposible probar el efecto interactivo de variables. Ahora hay razón para pensar que ciertos problemas no pueden resolverse a menos que se les aborde de forma multivariada. Un ejemplo de esto es la relación entre los métodos de enseñanza con el aprovechamiento escolar y otras variables. Es probable que los métodos de enseñanza, *per se*, no difieran mucho si sólo se estudian sus efectos simples. Los métodos de enseñanza trabajan en forma diferente bajo distintas condiciones, con diversos maestros y con alumnos variados. Se dice que los métodos "interactúan" con las condiciones y características de los docentes y de los estudiantes. Simon (1987) presentó otro ejemplo: un estudio de investigación sobre entrenamiento de pilotos propuesto por Williams y Adelson en 1954 no se podía realizar usando los métodos tradicionales de experimentación. El estudio proponía examinar 34 variables y su influencia en el entrenamiento de los pilotos. Con el uso de métodos tradicionales de investigación el número de variables bajo estudio era abrumador. Alrededor de 20 años después, Simon (1976) y Simon y Roscoe (1984) demostraron la forma en que se podía abordar con efectividad tales estudios usando diseños económicos de megafactor. Un ejemplo puede ayudarnos a entender este paso de razonamiento-deducción.

Suponga que un investigador está intrigado por la conducta agresiva. El investigador se pregunta por qué las personas son con frecuencia agresivas en situaciones donde este comportamiento puede ser inapropiado. La observación personal nos lleva a suponer que la conducta agresiva parece ocurrir cuando las personas han experimentado dificultades de uno u otro tipo. (Nótese la vaguedad del problema en este punto). Después de pensar por un tiempo, revisar la literatura para obtener algunas claves y llevar a cabo más observaciones, se formula la hipótesis: la frustración conduce a la agresividad. *La frustración* se define como el obstáculo para alcanzar una meta y la *agresividad* como la conducta caracterizada por ataque físico o verbal a otras personas u objetos.

Lo que sigue es una declaración como ésta: si la frustración conduce a la agresividad, entonces deberíamos encontrar un alto grado de agresividad entre los niños de escuelas restrictivas, que no les permiten mucha libertad ni posibilidad de expresión. De forma similar, en situaciones sociales difíciles, suponiendo que son frustrantes, esperaríamos más agresividad de lo "común". Si seguimos el razonamiento, si les diéramos a sujetos experimentales problemas interesantes para resolver, y después evitáramos que los solucionaran, podemos predecir algún tipo de conducta agresiva. En pocas palabras, el proceso de trasladarnos de un contexto amplio a una situación más específica se llama *razonamiento deductivo*.

El razonamiento puede, como se indicó antes, cambiar el problema. Podemos comprender que el problema inicial era sólo un caso especial de un problema más amplio, fundamental e importante. Podemos por ejemplo, iniciar con una hipótesis más limitada: las situaciones escolares restrictivas conducen a negativismo en los niños. Después podemos generalizar el problema a: la frustración induce la agresividad. Aun cuando es una forma diferente de pensamiento de lo arriba discutido, es importante, por lo que casi podríamos llamar su calidad heurística. El razonamiento puede ayudar a dirigirnos hacia

problemas más amplios, más básicos y por tanto más significativos, así como a aportar implicaciones operacionales (susceptibles de ser probadas) de la hipótesis original. Este tipo se llama *razonamiento inductivo*. Parte de un hecho particular hacia un enunciado general o hipótesis. Si uno no es cuidadoso este método puede inducir un razonamiento deficiente debido a su tendencia natural de excluir datos que no se ajustan a la hipótesis. El método de razonamiento inductivo tiende a buscar más datos de apoyo que a refutar evidencias.

Considere el estudio clásico de Peter Wason (Wason y Johnson-Laird, 1972) que ha suscitado un gran interés (Hoch 1986; Klayman y Ha, 1987). En este estudio se le pidió a estudiantes descubrir la regla que el experimentador tenía en mente al generar una secuencia de números. Un ejemplo fue generar una regla de la siguiente serie: "3, 5, 7". Se les dijo a los alumnos que podrían preguntar acerca de otras secuencias y que recibirían retroalimentación en cada serie sobre si se ajustaron o no a la regla pensada por el experimentador. Cuando los estudiantes se sintieran seguros podrían externar la regla. Algunos estudiantes indicaron: "9, 11, 13", y les dijeron que esta secuencia se ajustaba a la regla. Después siguieron con "15, 17, 19" y otra vez les respondieron que la serie correspondía. Los estudiantes entonces presentaron su respuesta: "la regla es tres números nones consecutivos", pero se les dijo que ésta *no* era la regla. Después de varias secuencias más propusieron contestaciones tales como: "números con incrementos de dos en dos", o bien "números nones con incrementos de valor dos". En cada uno de los casos se les indicó que ésa no era la regla que el experimentador había pensado. La regla que el experimentador tenía en mente era: "tres números positivos crecientes cualesquiera". Si los estudiantes hubieran propuesto las secuencias "8, 9, 10" o "1, 15, 4500" les habrían dicho que estos números también se ajustaban a la regla. Donde los estudiantes cometieron el error fue en probar sólo los casos que se ajustaban a la primera secuencia propuesta y que confirmaba su hipótesis.

Aunque simplificado en exceso, el estudio de Wason demostró lo que puede pasar en la investigación científica real. Un científico puede fácilmente condenarse a repetir el mismo tipo de experimento que siempre apoye la hipótesis.

Observación-prueba-experimento

A estas alturas debe quedar claro que la fase de observación-prueba-experimento forma sólo parte de todo el proceso científico. Si el problema ha sido bien planteado, la o las hipótesis se han formulado de manera adecuada y las implicaciones de las hipótesis se han deducido con cuidado, se puede presumir en este paso que el investigador es competente desde el punto de vista técnico.

La esencia de la prueba de hipótesis consiste en probar la *relación* expresada por la hipótesis. No se prueban las variables como tales, sino la relación entre ellas. La observación, la prueba y la experimentación tienen un propósito fundamental: probar empíricamente la relación del problema. Probar sin saber —al menos en alguna medida— *qué* y *por qué* uno evalúa implica un disparate. El hecho de sólo plantear un problema vago, por ejemplo, "¿cómo afecta al aprendizaje la educación abierta?" para después evaluar alumnos en las escuelas que dicen ser diferentes en cuanto a su "nivel de apertura", o preguntarnos: "¿cuáles son los efectos de la disonancia cognitiva?" para luego crear disonancia a través de manipulaciones experimentales y buscar los supuestos efectos, todo esto sólo podría llevarnos a información cuestionable. De forma similar, decir que se estudiarán los procesos de atribución sin realmente saber lo que se hace, o sin establecer relaciones entre las variables, constituye investigación sin sentido.

Otro aspecto importante es que por lo general no se prueba la hipótesis de manera directa. Como se indicó en el paso previo sobre el razonamiento, probamos las implicaciones que deducimos de las hipótesis. Nuestra hipótesis a probar puede ser: "los sujetos a quienes se les instruye para que eviten pensamientos indeseados estarán más preocupados con ellos que aquellos a los que se da una distracción". Ésta se dedujo de una hipótesis más amplia y general: "Cuantos mayores sean los esfuerzos para suprimir una idea, mayor será la preocupación sobre esta idea." No probamos "la supresión de ideas" o "la preocupación" sino la relación entre ellas, en este caso, la relación entre supresión de pensamientos indeseados y el nivel de preocupación (véase Wegner, Schneider, Carter y White, 1987; Wegner, 1989).

Dewey enfatizó que la secuencia temporal del pensamiento o cuestionamiento reflexivos no está fija. Podemos repetir y enfatizar lo que dijo en nuestro propio marco: los pasos del enfoque científico no están ordenados de manera impecable. El primer paso no se completa perfectamente antes de iniciar el segundo. Más aún, podemos hacer pruebas antes de deducir de forma adecuada las implicaciones de la hipótesis. La hipótesis, en sí misma, puede necesitar una mayor elaboración o refinamiento como resultado de las implicaciones deducidas de ella. Con frecuencia veremos que las hipótesis y su expresión parecen inadecuadas una vez que se deducen implicaciones a partir de ellas. Cuando una hipótesis es muy vaga se presenta la dificultad común de observar que una deducción es tan buena como otra; esto es, la hipótesis puede no conducir a pruebas precisas.

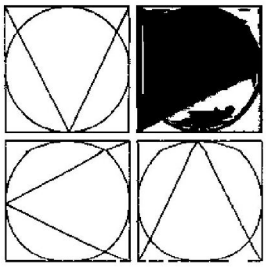
Retroalimentar el problema, la hipótesis y finalmente, la teoría de los resultados de la investigación es de la mayor importancia. Los teóricos e investigadores del aprendizaje, por ejemplo, con frecuencia han modificado sus teorías e investigaciones como resultado de hallazgos experimentales (véase Malone, 1991; Schunk, 1996; Hergenhahn, 1996). Teóricos e investigadores han estudiado los efectos del entorno y del entrenamiento tempranos en el desarrollo posterior. Kagan y Zentner (1996) revisaron los resultados de 70 estudios sobre las relaciones entre las experiencias de la edad temprana y la psicopatología en la vida adulta. Encontraron que es posible predecir la delincuencia juvenil a partir de la cantidad de impulsividad detectada en la etapa preescolar. Lynch, Short y Chua (1995) hallaron que el procesamiento musical estaba influido por la estimulación perceptual experimentada por el niño entre los seis meses y el año. Éstas y otras investigaciones han generado evidencia variada que converge en este problema de extrema importancia teórica y práctica. Una parte esencial de la investigación científica es el esfuerzo constante por replicar y verificar los hallazgos, por corregir la teoría con base en la evidencia empírica y por encontrar mejores explicaciones para los fenómenos naturales. Se puede incluso aseverar que la ciencia tiene un aspecto cíclico. Un investigador encuentra, por ejemplo, que *A* está relacionada con *B* de tal o cual forma. Entonces se conduce más investigación para determinar bajo qué otras condiciones *A* está relacionada de forma similar a *B*. Otros investigadores desafían esta teoría e investigación y ofrecen sus propias evidencias y explicaciones. El investigador original, se espera, modificará su trabajo a la luz de los nuevos datos: el proceso nunca termina.

Resumamos el llamado enfoque científico de la investigación. En primera instancia existe una duda, una barrera, una situación indeterminada que debe determinarse. El científico experimenta dudas vagas, disturbios emocionales e ideas incipientes. Hay un esfuerzo por formular el problema aunque sea de forma inadecuada. El científico entonces revisa la literatura y busca en su propia experiencia y en la de otros. Es frecuente que el investigador tenga que esperar un momento de inventiva: puede ser que ocurra, puede ser que no. Una vez que se cuenta con el problema formulado, con la o las preguntas básicas expresadas de manera apropiada, el resto es más fácil. Entonces, la hipótesis se construye, después de lo cual se deducen las implicaciones empíricas. Durante este proceso el proble-

ma original, y por supuesto la hipótesis original, pueden cambiarse, hacerse más generales o reducirse. Incluso pueden abandonarse. Más tarde se prueba la relación expresada por la hipótesis por medio de la observación y la experimentación. Con base en la evidencia procedente de la investigación, la hipótesis se apoya o rechaza. Esta información después retroalimenta al problema original, que se conserva o modifica de acuerdo con la evidencia. Dewey señaló que una fase del proceso puede expandirse y ser de gran importancia, otra puede reducirse, y puede haber más o menos pasos involucrados. La investigación rara vez es un asunto ordenado. De hecho, resulta mucho más desordenado que lo que la discusión anterior pueda sugerir. El orden y el desorden, sin embargo, no son de primera importancia. Lo que *sí* es importante es la racionalidad controlada de la investigación científica como un proceso de indagación reflexiva, la naturaleza interdependiente de las partes del proceso y la suprema importancia del problema y su enunciado.

RESUMEN DEL CAPÍTULO

1. Para comprender la compleja conducta humana es necesario entender el lenguaje y enfoque científicos.
2. La ciencia es una extensión sistemática y controlada del sentido común. Hay cinco diferencias entre el sentido común y la ciencia:
   a) La ciencia utiliza esquemas conceptuales y estructuras teóricas.
   b) La ciencia prueba de forma sistemática y empírica las teorías e hipótesis.
   c) La ciencia intenta controlar posibles causas extrañas.
   d) La ciencia busca relaciones de manera consciente y sistemática.
   e) La ciencia excluye explicaciones metafísicas (no demostrables).
3. Los cuatro métodos del conocimiento de Peirce son:
   a) Método de la tenacidad —influenciado por las creencias pasadas ya establecidas—.
   b) Método de autoridad —determinado por el peso de la tradición o la sanción pública—.
   c) Método *a priori* (también conocido como método de la intuición) —una natural inclinación hacia la verdad—.
   d) Método de la ciencia —autocorrectivo; los conceptos son objetivos y susceptibles de ser probados—.
4. Los estereotipos de la ciencia han limitado la comprensión de esta actividad por parte del público.
5. Visión y función de la ciencia
   a) Una visión estática considera a la ciencia como proveedora de información para el mundo; la ciencia aporta al cuerpo de información y al estado actual del conocimiento.
   b) La visión dinámica está interesada en la actividad de la ciencia (lo que hacen los científicos). Con ello aparece la visión heurística de la ciencia, que tiene un carácter de autodescubrimiento. La ciencia asume riesgos y resuelve problemas.
6. Los objetivos de la ciencia son:
   a) Generar teoría y explicar los fenómenos naturales.
   b) Promover la comprensión y desarrollar predicciones.
7. Una teoría tiene tres características:
   a) Posee un conjunto de propiedades con base en constructos definidos e interrelacionados.



CAPÍTULO 2

PROBLEMAS E HIPÓTESIS

- PROBLEMAS
- CRITERIOS DE LOS PROBLEMAS Y ENUNCIADOS DE PROBLEMAS
- HIPÓTESIS
- IMPORTANCIA DE LOS PROBLEMAS E HIPÓTESIS
- VIRTUDES DE LOS PROBLEMAS E HIPÓTESIS
- PROBLEMAS, VALORES Y DEFINICIONES
- GENERALIDAD Y ESPECIFICIDAD DE LOS PROBLEMAS E HIPÓTESIS
- LA NATURALEZA MULTIVARIABLE DE LA INVESTIGACIÓN Y PROBLEMAS DEL COMPORTAMIENTO
- COMENTARIOS FINALES: EL PODER ESPECIAL DE LAS HIPÓTESIS

Mucha gente cree que la ciencia es en lo fundamental una actividad de recolección de hechos. M.R. Cohen (1956/1997, p. 148) lo planteó de otra forma:

> No hay un progreso genuino en el discernimiento científico a través del método baconiano de acumular hechos empíricos sin una hipótesis o anticipación de la naturaleza. Sin alguna idea que nos guíe no sabemos qué hechos recolectar... no podemos determinar qué es y qué no es relevante.

Las personas sin formación científica con frecuencia tienen la idea que el científico es un individuo objetivo en extremo que recolecta datos sin tener ideas preconcebidas. Poincare (1952/1996, p. 143) señaló lo equivocado de esta idea: "Se dice a menudo que los experimentos deben realizarse sin ideas preconcebidas. Eso es imposible. No sólo haría que todo experimento fuera improductivo, sino que, aun deseándolo, no podríamos hacerlos."

Problemas

No siempre le es posible al investigador definir el problema de una manera simple, clara y completa. A menudo, puede tener una noción general, difusa e, inclusive, confusa del pro-

blema. Esto se debe a la naturaleza compleja de la investigación científica. Es posible incluso que, pueda tomarle años de exploración, reflexión e investigación el poder definir una pregunta de forma clara. Sin embargo, enunciar de manera adecuada el problema de investigación es una de las partes fundamentales del proceso. La dificultad para enunciar un problema de investigación de forma satisfactoria en un momento dado no debe hacernos perder de vista lo necesario y deseable que resulta.

Con esta dificultad en mente, podemos establecer un principio fundamental: si queremos resolver un problema, en general debemos conocerlo. Se puede decir que gran parte de la solución estriba en conocer lo que se trata de hacer. Otra parte está en entender qué es un problema y, en especial, un problema científico.

¿Qué constituye un buen enunciado del problema? Aunque los problemas de investigación difieren en gran medida y no existe una fórmula "correcta" para enunciar problemas, es posible aprender y utilizar para nuestro beneficio ciertas características de los problemas y de los enunciados de problemas. Para empezar, consideremos dos o tres ejemplos de problemas de investigación publicados y estudiemos sus características. Primero, tomemos el problema del estudio realizado por Hurlock (1925),[1] mencionado en el capítulo 1: ¿Cuáles son los efectos de diferentes tipos de incentivos en el rendimiento del alumno? Observe que el problema está enunciado en forma de pregunta. En este campo, la forma más simple es la mejor. También conviene señalar que el problema establece una relación entre variables, en este caso, entre las variables *incentivos* y *rendimiento del alumno* (logro). (El término *variable* será definido de manera formal en el capítulo 3. Por ahora, se usará para nombrar un fenómeno o un constructo, que asume un conjunto de diferentes valores numéricos.)

Un *problema*, entonces, es un enunciado u oración interrogativa que pregunta: ¿qué relación existe entre dos o más variables? La respuesta constituye aquello que se busca en la investigación. Un problema, en la mayoría de los casos, tendrá dos o más variables. En el ejemplo de Hurlock, el enunciado del problema relaciona incentivos con rendimiento del alumno. Otro problema, estudiado en el experimento clásico de Bahrick (1984, 1992) está asociado con preguntas de edad y vejez: ¿Cuánto de lo que ahora estás estudiando recordarás dentro de diez años? ¿Cuánto de esto mismo podrás recordar dentro de cincuenta años? ¿Cuánto recordarás después, si nunca lo utilizas? La pregunta formal de Bahrick es: ¿la memoria semántica involucra procesos separados? Una variable es la cantidad de tiempo que transcurre desde que el material se aprendió por primera vez; la segunda podría ser la calidad del aprendizaje original; y la otra variable es el recuerdo (u olvido). Veamos otro problema de Little, Sterling y Tingstrom (1996), que es muy diferente: ¿Influyen las claves geográficas y las características raciales en la atribución (culpa percibida)? Una variable son las claves geográficas; la segunda sería la información racial, y la tercera, la atribución.

No todos los problemas de investigación contienen dos o más variables claras. Por ejemplo, en psicología experimental, el foco de la investigación con frecuencia está en procesos psicológicos como la memoria y la categorización. Rosch (1973) en su influyente y justificadamente bien conocido estudio de categorías perceptuales hizo la siguiente pregunta: ¿Existen categorías no arbitrarias ("naturales") de color y forma? Aunque la relación entre dos o más variables no es aparente en este enunciado del problema, en la investigación real las categorías estaban relacionadas con el aprendizaje. Hacia el final de este libro se verá que los problemas de investigación factorial analítica también carecen

[1] Cuando referimos problemas e hipótesis de la literatura, no siempre usamos las palabras de los autores. De hecho, los enunciados de muchos de los problemas son nuestros y no de los autores citados. Algunos autores sólo usan enunciados de problemas, algunos sólo hipótesis, y otros usan ambos.

de la forma de las relaciones antes planteada. Sin embargo, en la mayoría de los problemas de investigación del comportamiento, se estudian las relaciones entre dos o más variables; por ello, enfatizaremos ese tipo de enunciados de relación.

Criterios de los problemas y enunciados de problemas

Existen tres criterios de buenos problemas y enunciados de problemas. El primero: el problema debe expresar una relación entre dos o más variables. En efecto plantea preguntas como la siguiente: ¿*A* está relacionada con *B*? ¿Cómo están relacionadas *A* y *B* con *C*? ¿Cómo está *A* relacionada con *B* bajo las condiciones *C* y *D*? La excepción a esta consideración ocurre casi siempre en investigación metodológica o taxonómica.

Segundo: el problema debe ser enunciado de manera clara y sin ambigüedades en forma de pregunta. En lugar de decir, por ejemplo "el problema es..." o "El propósito de este estudio es..." resulta necesario plantear una pregunta. Las preguntas tienen la virtud de presentar los problemas directamente. El propósito de un estudio no es por fuerza el mismo que el problema de un estudio. El propósito del estudio de Hurlock, por ejemplo, fue arrojar luz sobre el uso de incentivos en las situaciones escolares. El problema consistió en la pregunta acerca de la relación entre incentivos y rendimiento. Otra vez, mientras más simple, mejor: formule una pregunta.

El tercer criterio con frecuencia es difícil de satisfacer. Demanda que el problema y su enunciado *impliquen* la posibilidad de ser sometidos a una prueba empírica. Un problema que no contenga implicaciones para probar las relaciones que enuncia, no constituye un problema científico. Esto significa no sólo que se enuncie una relación real, sino también que las variables de la relación puedan ser medidas de alguna forma. Hay muchas preguntas interesantes e importantes que no constituyen preguntas científicas tan sólo porque no son susceptibles de prueba. Ciertas preguntas filosóficas y teológicas, aunque importantes para quienes las consideran, no pueden ser probadas empíricamente, por lo que no generan interés para el científico como tal. La pregunta epistemológica "¿Cómo conocemos?" es una pregunta de ese tipo. La educación plantea muchas preguntas interesantes pero no científicas, por ejemplo "¿Mejora la educación democrática el aprendizaje de los jóvenes?" "¿Son buenos los procesos grupales para los niños?" Estas preguntas pueden ser etiquetadas como metafísicas en el sentido en que están, al menos así enunciadas, fuera de la posibilidad de una prueba empírica. Las principales dificultades estriban en que algunas no constituyen relaciones, y es muy difícil o imposible definir la mayoría de sus constructos de forma que puedan ser medidos.

Hipótesis

Una *hipótesis* es un enunciado conjetural de la relación entre dos o más variables. Las hipótesis siempre se presentan en forma de enunciados declarativos y relacionan, de manera general o específica, variables con variables. Hay dos criterios que definen a las "buenas" hipótesis y a sus enunciados. Son los mismos que mencionamos para los problemas y sus enunciados. 1) Las hipótesis son enunciados acerca de las relaciones entre variables. 2) Las hipótesis contienen implicaciones claras para probar las relaciones enunciadas. Estos criterios significan que los enunciados de hipótesis contienen dos o más variables, que son medibles o pueden serlo, y que especifican cómo están relacionadas las variables.

Permítanos mencionar tres hipótesis de la literatura y aplicarles estos criterios. La primera hipótesis procede de un estudio de Wegner y colaboradores. (1987) que parece desafiar el sentido común: a mayor supresión de pensamientos indeseados, mayor preocu-

pación por ellos (represión ahora; obsesión más tarde). Aquí se establece una relación entre una variable, supresión de una idea o pensamiento, y otra variable, preocupación u obsesión. Dado que ambas se definen y miden con facilidad, las implicaciones para probar la hipótesis también se conciben sin esfuerzo. Los criterios están satisfechos. En el estudio de Wegner y colaboradores, se les pidió a los sujetos que *no* pensaran en un "oso blanco". Cada vez que pensaran en él, debían de tocar una campana. El número de campanadas indicaba el nivel de preocupación. Una segunda hipótesis, que resulta inusual y corresponde al estudio de Ayres y Hughes (1986), enuncia la relación de una forma que llamamos nula: el nivel de ruido o música no tiene efecto en el funcionamiento visual. La relación se establece con claridad: una variable, intensidad del sonido (por ejemplo, música), se relaciona con otra, funcionamiento visual, a través de las palabras "no tiene efecto en". En el criterio de potencia de ser probada, sin embargo, encontramos dificultades. Nos enfrentamos con el problema de definir "funcionamiento visual" e "intensidad" de forma que puedan medirse. Si podemos resolver este problema de manera satisfactoria, entonces, tenemos en definitiva una hipótesis. Ayres y Hughes lo hicieron al definir intensidad como 107 decibeles y funcionamiento visual en términos de una puntuación en una tarea de agudeza visual. Esta hipótesis permitió contestar una pregunta que la gente con mucha frecuencia se hace: "¿por qué bajamos el volumen del estéreo del auto cuando *buscamos* una dirección?". Ayres y Hughes encontraron una caída marcada en el funcionamiento perceptual cuando el nivel de música llegaba a 107 decibeles.

La tercera hipótesis representa una categoría numerosa e importante. En ella la relación es indirecta, oculta. En general enuncia que los grupos *A* y *B* diferirán en alguna característica. Por ejemplo: las mujeres creen, con mayor frecuencia que los hombres, que deben perder peso aun cuando éste se encuentre dentro de los límites normales (Fallon y Rozin, 1985). Esto es, que las mujeres difieren de los hombres en cuanto a la percepción de su figura corporal. Observe que este enunciado está un paso más allá de la hipótesis real que puede plantearse como: la percepción de la figura corporal es, en parte, una función del género. Si los enunciados posteriores constituyeran la hipótesis enunciada, entonces la primera podría llamarse una subhipótesis o una predicción específica basada en la hipótesis original.

Consideremos otra hipótesis de este tipo pero dando un paso más adelante. Los individuos que tienen características iguales o similares tendrán actitudes similares hacia objetos cognitivos significativamente relacionados con su papel ocupacional (Saal y Moore, 1993). (*Los objetos cognitivos* se definen como algo concreto o abstracto, percibido y "conocido" por los individuos. Personas, grupos, ascenso en el trabajo o en las calificaciones, el gobierno y la educación son algunos ejemplos.) La relación en este caso es, por supuesto, entre características personales y actitudes (hacia un objeto cognitivo relacionado con la característica personal, por ejemplo, género y actitudes hacia otros que reciben una promoción). Para probar esta hipótesis, sería necesario tener al menos dos grupos, cada uno con una característica diferente, y después comparar las actitudes de ambos grupos. Por ejemplo, como en el caso del estudio de Saal y Moore, la comparación sería entre hombres y mujeres. Serían comparados en relación a su evaluación hacia el ascenso dado a un compañero de trabajo del mismo sexo o del opuesto. En este ejemplo, se satisfacen los criterios.

Importancia de los problemas e hipótesis

Hay poca duda de que las hipótesis son herramientas importantes e indispensables de la investigación científica. Existen tres razones principales para esta creencia. La primera es que son, digamos, los instrumentos de trabajo de la teoría. Las hipótesis pueden deducirse

a partir de la teoría y de otras hipótesis. Si por ejemplo, trabajamos en una teoría sobre la agresividad, se presume que buscamos causas y efectos del comportamiento agresivo. Es posible que hayamos observado casos de comportamiento agresivo ocurrido después de circunstancias frustrantes. La teoría, entonces, puede incluir la proposición: la frustración produce agresividad (Berkowitz, 1983; Dill y Anderson, 1995; Dollard, Doob, Miller, Mowrer, y Sears, 1939). A partir de esta amplia hipótesis, podemos deducir hipótesis más específicas, como por ejemplo: impedir que los niños alcancen sus metas deseadas (frustración) generará pleitos entre ellos (agresión); si los niños son privados del amor paterno (frustración), reaccionarán en parte, con un comportamiento agresivo.

La segunda razón es que es posible someter a prueba las hipótesis y demostrar que son probablemente verdaderas o probablemente falsas. No se prueban hechos aislados, como se dijo antes, sólo relaciones. Es probable que la principal razón de usar hipótesis en la investigación científica sea que constituyen proposiciones relacionales. En esencia, son predicciones del tipo: "si *A*, entonces *B*", que utilizamos para probar la relación entre *A* y *B*. Dejamos que los hechos tengan la oportunidad de establecer la probable veracidad o falsedad de la hipótesis.

La tercera razón es que las hipótesis son herramientas poderosas para el avance del conocimiento porque permiten al científico ir más allá de sí mismo. Aunque desarrolladas por humanos, las hipótesis existen, pueden ser probadas y puede demostrarse que son probablemente correctas o incorrectas de manera independiente a los valores y opiniones de una persona (sesgos). Esto resulta crítico: no habría ciencia, en sentido completo alguno, sin las hipótesis.

Tan importantes como las hipótesis son los problemas tras ellas. Como Dewey (1938/1982, pp. 105-107) ha señalado, la investigación por lo general empieza con un problema. Indica que primero hay una situación indeterminada en la que las ideas son vagas, aparecen dudas, y el pensador queda perplejo. Agrega que el problema no se enuncia; de hecho, no puede ser enunciado hasta que uno ha experimentado una situación tan indeterminante.

La indeterminación, sin embargo, deberá, en última instancia ser eliminada. Aunque es cierto, como se señaló antes, que el investigador con frecuencia puede tener sólo una noción general y difusa del problema, tarde o temprano habrá de definir una idea clara de lo que el problema es. Aunque este enunciado parezca obvio, una de las cosas más difíciles de lograr, es enunciar de una manera clara y completa el problema de investigación. En otras palabras, uno debe saber qué es lo que trata encontrar. Cuando por fin se identifica, el problema ya está en camino a la solución.

Virtudes de los problemas e hipótesis

Los problemas y las hipótesis tienen virtudes importantes: 1) dirigen la investigación (las relaciones expresadas en las hipótesis indican al investigador lo que debe hacer); 2) los problemas e hipótesis, dado que son de ordinario enunciados relacionales generalizados, permiten al investigador deducir manifestaciones empíricas específicas implicadas en ellos. Podemos decir, de acuerdo con Guida y Ludlow (1989): si es un hecho verdadero que los niños de un tipo de cultura (Chile) tienen un mayor grado de ansiedad que los niños de otro tipo de cultura (blancos estadounidenses), entonces los niños en la cultura chilena deben rendir menos en lo académico que los niños en la cultura estadounidense. Los niños chilenos quizá también debieran presentar una menor autoestima o un *locus* de control más externo en lo que se refiere a la escuela y a la labor académica.

Hay diferencias importantes entre problemas e hipótesis. Las hipótesis, si están enunciadas de manera apropiada, pueden ser probadas. Una hipótesis dada puede ser demasia-

do amplia para ser probada de forma directa; sin embargo, si es una "buena" hipótesis, es posible deducir a partir de ella otras que sí lo sean. Los hechos o las variables no se prueban como tales. Se prueban las relaciones enunciadas por las hipótesis. Un problema no puede ser resuelto de manera científica a menos que se reduzca a su forma de hipótesis, ya que un problema es una pregunta, generalmente de naturaleza amplia, que no puede probarse en forma directa. No se someten a prueba preguntas como: ¿la presencia o ausencia de otra persona en un sanitario público afecta la higiene personal? (Pedersen, Keithly y Brady, 1986). ¿Las sesiones de consejería grupal disminuyen el nivel de morbilidad psiquiátrica en oficiales de policía? (Doctor, Cutris e Issacs, 1994). Quizás uno probaría una o más hipótesis deducidas de estas preguntas. Por ejemplo, para estudiar el último problema, uno puede hipotetizar que los oficiales de policía que asisten a sesiones de consejería para reducir el estrés requerirán menos días de incapacidad por enfermedad que aquéllos que no asisten. La hipótesis para el primer problema podría indicar que la presencia de una persona en un sanitario público hará que otras se laven las manos.

Los problemas e hipótesis permiten que avance el conocimiento científico al ayudar al investigador a confirmar o refutar una teoría. Suponga que un investigador en el campo de la psicología aplica a unos sujetos tres o cuatro pruebas, entre las cuales hay una para evaluar la ansiedad relacionada con una prueba aritmética. Al calcular de manera rutinaria las correlaciones entre las tres o cuatro pruebas, uno encuentra que la correlación entre ansiedad y aritmética es negativa. De lo anterior se deduce que a mayor ansiedad, menor puntuación en la prueba de aritmética. Sin embargo, es muy probable que esta relación sea fortuita e incluso espuria, pero si el investigador hubiera hipotetizado la relación con base en una teoría, tendría mayor confianza en sus resultados. El investigador que no hipotetiza relaciones en forma previa, no permite que los hechos prueben o rechacen nada. Las palabras *probar* y *rechazar* no deben tomarse en su sentido literal: una hipótesis nunca se prueba o refuta realmente. Para ser más precisos, deberíamos decir algo del tipo de: el peso de la evidencia está del lado de la hipótesis o el peso de la evidencia arroja dudas sobre la hipótesis. Braithwaite (1953/1996, p. 14) dice:

> De este modo, la evidencia empírica nunca prueba la hipótesis: en casos apropiados podemos decir que *se establece* (itálicas agregadas) la hipótesis, lo que significa que la evidencia hace que sea razonable aceptar la hipótesis; pero ésta nunca *prueba* la hipótesis en el sentido de que la hipótesis sea una consecuencia lógica de la evidencia.

Este uso de la hipótesis es similar a participar en un juego de azar. Se establecen las reglas del juego y se definen las apuestas por adelantado. Uno no puede cambiar las reglas después de un resultado, como tampoco se pueden cambiar las apuestas una vez hechas. Uno no tira los dados primero y luego apuesta. No sería "justo". De igual forma, si en primera instancia se recolectan datos y después se toma uno de ellos y se llega a una conclusión con base en él, se han violado las reglas del juego científico. El juego no es "justo" porque el investigador puede capitalizar fácilmente, digamos, dos relaciones importantes de las cinco a prueba. Las otras tres, por lo general, se olvidan. En un juego "justo", se cuenta cada tiro del dado, en el sentido de que se gana o no con base en el resultado de cada tirada.

Las hipótesis dirigen la investigación. Como Darwin señaló hace más de 100 años, todas las observaciones han de ser a favor o en contra de algún punto de vista para tener alguna utilidad. Las hipótesis incorporan aspectos de la teoría bajo prueba de forma susceptible o casi susceptible de ser probada. Antes se dio un ejemplo de la teoría del reforzamiento en el que se dedujeron hipótesis demostrables a partir del problema general. Podemos demostrar la importancia del reconocimiento de esta función de las hipótesis al introducirnos por la puerta trasera y usar una teoría muy difícil o quizás imposible de probar. La teoría de Freud de la ansiedad incluye el constructo de la represión. Con este

término Freud se refería a la introducción forzada de ideas inaceptables en lo profundo del inconsciente. Para probar la teoría freudiana de la ansiedad es necesario deducir relaciones sugeridas por la teoría. Estas deducciones por fuerza deberán incluir el concepto de represión que implica el constructo del inconsciente. Es posible formular hipótesis que utilizan estos constructos; para probar la teoría han de ser formuladas de esta manera. Pero probarlos resulta más difícil debido a la extrema dificultad para definir términos como "represión" e "inconsciente" de manera que puedan medirse. Hasta hoy nadie ha tenido éxito al definir estos dos constructos sin apartarse en gran medida del significado y uso freudianos originales. Las hipótesis constituyen, entonces, puentes importantes entre la teoría y la investigación empírica.

Problemas, valores y definiciones

Para establecer con claridad la naturaleza de los problemas y de las hipótesis, analizaremos ahora dos o tres errores comunes. En primer término, los problemas científicos no son preguntas morales ni éticas: ¿Son las medidas disciplinarias de tipo punitivo perjudiciales para los niños? ¿Debiera ser el liderazgo de una organización de tipo democrático? ¿Cuál es la mejor forma de enseñar a los estudiantes universitarios? Formular estas preguntas equivale a presentar cuestionamientos de valor y juicio que la ciencia no puede contestar. Muchas de las que se han llamado hipótesis no lo son en absoluto. Por ejemplo: el método de enseñanza a pequeños grupos es mejor que el método expositivo. Éste es un enunciado de valor; constituye un acto de fe, no una hipótesis. Si fuera posible establecer una relación entre las variables, y definirlas de manera que se pudiera probar esa relación, entonces podríamos contar con una hipótesis. Pero no hay una forma científica de someter a prueba preguntas de valor.

Una forma rápida y relativamente fácil de detectar preguntas y enunciados de valor consiste en buscar palabras como *debe, debería, mejor que* (en lugar de *mayor que*), así como palabras similares que indiquen juicios culturales o personales, o preferencias (sesgos). Sin embargo, los enunciados de valor son engañosos. Aunque resulta obvio que un enunciado que incluye la palabra "debería" es un enunciado de valor, otros tipos no son tan evidentes. Consideremos el enunciado: los métodos autoritarios de enseñanza conducen a un pobre aprendizaje. En este caso sí hay una relación. Pero el enunciado falla como una hipótesis científica en tanto que incorpora dos expresiones de valor: "métodos autoritarios de enseñanza" y "pobre aprendizaje", ninguna de las cuales puede definirse con propósitos de medición sin borrar las palabras *autoritario* y *pobre*.[2]

Con frecuencia se formula otra clase de enunciados que no constituyen hipótesis o que son hipótesis pobres, en especial en el campo de la educación. Consideremos, por ejemplo, el siguiente: los cursos de tronco común representan una experiencia enriquecedora. Otro tipo de enunciado que se usa con frecuencia, es la generalización vaga: es posible identificar las habilidades de lectura en el segundo grado; La meta del individuo auténtico es la autorrealización; El prejuicio se relaciona con ciertos rasgos de personalidad.

Otro defecto común de los enunciados de problema aparece a menudo en las tesis doctorales: enlistar aspectos metodológicos o "problemas" como subproblemas. Estos aspectos metodológicos poseen dos características que hacen fácil detectarlos: 1) No son

[2] Un caso ya casi clásico del uso de la palabra *autoritario* es la frase que a veces se escucha entre educadores: el método expositivo es autoritario. Esto parece indicar que quien lo dice no gusta del método expositivo y lo considera negativo. De forma similar, una de las formas más efectivas de criticar a un maestro es decir que es autoritario.

¿Qué es la ciencia?

1. Introducción

Mientras los animales inferiores sólo están en el mundo, el hombre trata de entenderlo: y, sobre la base de su inteligencia imperfecta pero perfectible del mundo, el hombre intenta enseñorearse de él para hacerlo más confortable. En este proceso, construye un mundo artificial: ese creciente cuerpo de ideas llamado "ciencia", que puede caracterizarse como conocimiento racional, sistemático, exacto, verificable y por consiguiente falible. Por medio de la investigación científica, el hombre ha alcanzado una reconstrucción conceptual del mundo que es cada vez más amplia, profunda y exacta.

Un mundo le es dado al hombre; su gloria no es soportar o despreciar este mundo, sino enriquecerlo construyendo otros universos. Amasa y remoldea la naturaleza sometiéndola a sus propias necesidades; construye la sociedad y es a su vez construido por ella; trata luego de remoldear este ambiente artificial para adaptarlo a sus propias necesidades animales y espirituales, así como a sus sueños: crea así el mundo de los artefactos y el mundo de la cultura. La ciencia como actividad —como investigación— pertenece a la vida social; en

cuanto se la aplica al mejoramiento de nuestro medio natural y artificial, a la invención y manufactura de bienes materiales y culturales, la ciencia se convierte en tecnología. Sin embargo, la ciencia se nos aparece como la más deslumbrante y asombrosa de las estrellas de la cultura cuando la consideramos como un bien por sí mismo, esto es, como un sistema de ideas establecidas provisionalmente (conocimiento científico), y como una actividad productora de nuevas ideas (investigación científica). Tratemos de caracterizar el conocimiento y la investigación científicos tal como se los conoce en la actualidad.

2. *Ciencia formal y ciencia fáctica*

No toda investigación científica procura el conocimiento objetivo. Así, la lógica y la matemática —esto es, los diversos sistemas de lógica formal y los diferentes capítulos de la matemática pura— son racionales, sistemáticos y verificables, pero no son objetivos, no nos dan informaciones acerca de la realidad: simplemente no se ocupan de los hechos. La lógica y la matemática tratan de entes ideales; estos entes, tanto los abstractos como los interpretados, sólo existen en la mente humana. A los lógicos y matemáticos no se les da objetos de estudio: ellos construyen sus propios objetos. Es verdad que a menudo lo hacen por abstracción de objetos reales (naturales y sociales), como es el caso de las figuras geométricas y de los números

enteros. Más aún, el trabajo del lógico o del matemático satisface a menudo las necesidades del naturalista, del sociólogo o del tecnólogo, y es por esto que la sociedad los tolera y, ahora, hasta los estimula. Pero la materia prima que emplean los lógicos y los matemáticos no es fáctica sino ideal.

Por ejemplo, el concepto de número abstracto nació, sin duda, de la coordinación (correspondencia biunívoca) de conjuntos de objetos materiales, tales como dedos, por una parte, y guijarros, por la otra; pero no por esto aquel concepto se reduce a esta operación manual, ni a los signos que se emplean para representarlo. Los números no existen fuera de nuestros cerebros. Los objetos materiales son numerables siempre que sean discontinuos; pero no son números; tampoco son números puros (abstractos) sus cualidades o relaciones. En el mundo real encontramos 3 libros, en el mundo de la ficción construimos 3 platos voladores. Pero ¿quién vio jamás un 3, un simple 3?

La lógica y la matemática, por ocuparse de inventar entes formales y de establecer relaciones entre ellos, se llaman a menudo *ciencias formales*, precisamente porque sus objetos no son cosas ni procesos sino, para emplear el lenguaje pictórico, formas en las que se puede verter un surtido ilimitado de contenidos, tanto fácticos como empíricos. Esto es, podemos establecer *correspondencias* entre esas formas (u objetos formales), por una parte, y cosas y procesos pertenecientes a cualquier nivel de la realidad, por la otra. Así es como

la física, la química, la fisiología, la psicología, la economía y las demás ciencias recurren a la matemática, empleándola como herramienta para representar con precisión las complejas relaciones que se encuentran entre los hechos, y entre los diversos aspectos de los hechos; dichas ciencias no identifican las formas ideales con los objetos concretos, sino que interpretan las primeras en términos de hechos y de experiencias (o, lo que es equivalente, formalizan enunciados fácticos).

Lo mismo vale para la lógica formal: algunas de sus partes —en particular, pero no exclusivamente, la lógica proposicional bivalente— pueden hacerse corresponder a aquellas entidades psíquicas que llamamos pensamientos. Semejante aplicación de las ciencias de la forma pura a la inteligencia del mundo de los hechos se efectúa asignando diferentes interpretaciones a los objetos formales. Estas interpretaciones son, dentro de ciertos límites, arbitrarias; vale decir, se justifican por el éxito, la conveniencia o la ignorancia. En otras palabras, el significado fáctico o empírico que se les asigna a los objetos formales no es una propiedad intrínseca de los mismos. De esta manera, las ciencias formales jamás entran en conflicto con la realidad. Esto explica la paradoja de que, siendo formales, se "aplican" a la realidad: en rigor no se aplican, sino que se emplean en la vida cotidiana y en las ciencias fácticas a condición de que se les superpongan reglas de correspondencia adecuada. En suma, la lógica y la matemática establecen contacto con la realidad a través del

puente del lenguaje, tanto el ordinario como el científico.

Tenemos así una primera gran división de las ciencias, en formales (o ideales) y fácticas (o materiales). Esta ramificación preliminar tiene en cuenta el *objeto* o tema de las respectivas disciplinas; también da cuenta de la diferencia de especie entre los enunciados que se proponen establecer las ciencias formales y las fácticas: mientras los enunciados formales consisten en relaciones entre signos, los enunciados de las ciencias fácticas se refieren, en su mayoría, a entes extracientíficos: a sucesos y procesos. Nuestra división también tiene en cuenta el *método* por el cual se ponen a prueba los enunciados verificables: mientras las ciencias formales se contentan con la lógica para demostrar rigurosamente sus teoremas (los que, sin embargo, pudieron haber sido adivinados por inducción común o de otras maneras), las ciencias fácticas necesitan más que la lógica formal: para confirmar sus conjeturas necesitan de la observación y/o el experimento. En otras palabras, las ciencias fácticas tienen que mirar las cosas y, siempre que les sea posible, deben procurar cambiarlas deliberadamente para intentar descubrir en qué medida sus hipótesis se adecuan a los hechos.

Cuando se demuestra un teorema lógico o matemático no se recurre a la experiencia: el conjunto de postulados, definiciones, reglas de formación de las expresiones dotadas de significado y reglas de inferencia deductiva —en suma la base de la teoría dada—, es necesario y suficiente para

ese propósito. La demostración de los teoremas no es sino una deducción: es una operación confinada a la esfera teórica, aun cuando a veces los teoremas mismos (no sus demostraciones) sean sugeridos en alguna esfera extramatemática, y aun cuando su prueba (pero no su primer descubrimiento) pueda realizarse con ayuda de calculadoras electrónicas. Por ejemplo, cualquier demostración rigurosa del teorema de Pitágoras prescinde de mediciones y emplea figuras sólo como ayuda psicológica al proceso deductivo. Que el teorema de Pitágoras haya sido el resultado de un largo proceso de inducción conectado a operaciones prácticas de mediciones de tierras, es objeto de la historia, la sociología y la psicología del conocimiento.

La matemática y la lógica son, en suma, ciencias deductivas. El proceso constructivo, en que la experiencia desempeña un gran papel sugestivo, se limita a la formación de los puntos de partida (axiomas). En matemática la verdad consiste, por esto, en la coherencia del enunciado dado con un sistema de ideas admitido previamente: por esto, la verdad matemática no es absoluta, sino relativa a ese sistema, en el sentido de que una proposición que es válida en una teoría puede dejar de ser lógicamente verdadera en otra teoría. (Por ejemplo, en el sistema de aritmética que empleamos para contar las horas del día, vale la proposición $24 + 1 = 1$.) Más aún las teorías matemáticas abstractas, esto es, que contienen términos no interpretados (signos a los que no se atribuye un

significado fijo, y que, por lo tanto, pueden adquirir distintos significados) pueden desarrollarse sin poner atención al problema de la verdad.

Considérese el siguiente axioma de cierta teoría abstracta (no interpretada): "Existe por lo menos un x tal que es F". Se puede dar un número ilimitado de interpretaciones (modelos) de este axioma, dándose a 'x' y 'F' otros tantos significados. Si decimos que 'F' designa punto, obtenemos un modelo geométrico dado; si adoptamos la convención de que 'F' designa número, obtenemos un cierto modelo aritmético, y así sucesivamente. En cuanto "llenamos" la forma vacía con un contenido específico (pero todavía matemático), obtenemos un sistema de entes lógicos que tienen el privilegio de ser verdaderos o falsos dentro del sistema dado de proposiciones: a partir de ahí tenemos que habérnoslas con el problema de la verdad matemática. Aun así tan sólo las conclusiones (teoremas) tendrán que ser verdaderas: los axiomas mismos pueden elegirse a voluntad. La batalla se habrá ganado si se respeta la coherencia lógica, esto es, si no se violan las leyes del sistema de lógica que se ha convenido usar.

En las ciencias fácticas, la situación es enteramente diferente. En primer lugar, ellas no emplean símbolos vacíos (variables lógicas), sino tan sólo símbolos interpretados; por ejemplo, no involucran expresiones tales como "x es F", que no son verdaderas ni falsas. En segundo lugar, la racionalidad —esto es, la coherencia con un sistema de ideas aceptado previamente— es necesaria pero

no suficiente para los enunciados fácticos; en particular, la sumisión a algún sistema de lógica es necesaria pero no es una garantía de que se obtenga la verdad. Además de la racionalidad, exigimos de los enunciados de las ciencias fácticas que sean *verificables en la experiencia*, sea indirectamente (en el caso de las hipótesis generales), sea directamente (en el caso de las consecuencias singulares de las hipótesis). Únicamente después que haya pasado las pruebas de la verificación empírica podrá considerarse que un enunciado es adecuado a su objeto, o sea, que es verdadero, y aun así hasta nuevo aviso. Por esto es que el conocimiento fáctico verificable se llama a menudo *ciencia empírica*.

En resumidas cuentas, la coherencia es necesaria pero no suficiente en el campo de las ciencias de hechos: para afirmar que un enunciado es (probablemente) verdadero se requieren datos empíricos (proposiciones acerca de observaciones o experimentos). En última instancia, sólo la experiencia puede decirnos si una hipótesis relativa a cierto grupo de hechos materiales es adecuada o no. El mejor fundamento de esta regla metodológica que acabamos de enunciar es que la experiencia le ha enseñado a la humanidad que el conocimiento de hecho no es convencional, que si se busca la comprensión y el control de los hechos debe partirse de la experiencia. Pero la experiencia no garantizará que la hipótesis en cuestión sea la única verdadera: sólo nos dirá que es *probablemente* adecuada, sin excluir por ello la posibilidad de que un estudio ulterior pueda dar mejores

aproximaciones en la reconstrucción conceptual del trozo de realidad escogido. El conocimiento fáctico, aunque racional, es esencialmente probable; dicho de otro modo: la inferencia científica es una red de inferencias deductivas (demostrativas) y plausibles (inconcluyentes).

Las ciencias formales demuestran o prueban; las ciencias fácticas verifican (confirman o disconfirman) hipótesis que en su mayoría son provisionales. La demostración es completa y final; la verificación es incompleta y por ello temporaria. La naturaleza misma del método científico impide la confirmación final de las hipótesis fácticas. En efecto, los científicos no sólo procuran acumular elementos de prueba de sus suposiciones multiplicando el número de casos en que ellas se cumplen; también tratan de obtener casos desfavorables a sus hipótesis, fundándose en el principio lógico de que una sola conclusión que no concuerde con los hechos tiene más peso que mil confirmaciones. Por ello, mientras las teorías formales pueden ser llevadas a un estado de perfección (o estancamiento), los sistemas teóricos relativos a los hechos son esencialmente defectuosos; cumplen, pues, la condición necesaria para ser perfectibles. En consecuencia, si el estudio de las ciencias formales puede vigorizar el hábito del rigor, el estudio de las ciencias fácticas puede inducirnos a considerar el mundo como inagotable, y al hombre como una empresa inconclusa e interminable.

Las diferencias de método, tipo de enunciados y referentes, que separan las ciencias fácticas de

las formales, impiden que se las examine conjuntamente más allá de cierto punto. Por ser una ficción seria, rigurosa y a menudo útil, pero ficción al cabo, la ciencia formal requiere un tratamiento especial. En lo que sigue nos concentraremos en la ciencia fáctica. Daremos un vistazo a las características peculiares de las ciencias de la naturaleza y de la sociedad en su estado actual, con la esperanza de que la ciencia futura enriquezca sus cualidades o, al menos, de que las civilizaciones por venir hagan mejor uso del conocimiento científico.

Los rasgos esenciales del tipo de conocimiento que alcanzan las ciencias de la naturaleza y de la sociedad son la *racionalidad* y la *objetividad*. Por conocimiento racional se entiende:

a) que está constituido por conceptos, juicios y raciocinios, y no por sensaciones, imágenes, pautas de conducta, etc. Sin duda, el científico percibe, forma imágenes (p. ej., modelos visualizables) y hace operaciones; pero tanto el punto de partida como el punto final de su trabajo son ideas;

b) que esas ideas pueden combinarse de acuerdo con algún conjunto de reglas lógicas, con el fin de producir nuevas ideas (inferencia deductiva). Éstas no son enteramente nuevas desde un punto de vista estrictamente lógico, puesto que están implicadas por las premisas de la deducción; pero son gnoseológicamente nuevas en la medida en que expresan conocimientos de los que no se tenía conciencia antes de efectuarse la deducción;

c) que esas ideas no se amontonan caóticamente o, simplemente, en forma cronológica, sino

que se organizan en sistemas de ideas, esto es, en conjuntos ordenados de proposiciones (teorías).

Que el conocimiento científico de la realidad es objetivo, significa:

a) que concuerda aproximadamente con su objeto; vale decir, que busca alcanzar la verdad fáctica;

b) que verifica la adaptación de las ideas a los hechos recurriendo a un comercio peculiar con los hechos (observación y experimento), intercambio que es controlable y hasta cierto punto reproducible.

Ambos rasgos de la ciencia fáctica, la racionalidad y la objetividad, están íntimamente soldados. Así, por ejemplo, lo que usualmente se verifica por medio del experimento es alguna consecuencia —extraída por vía deductiva— de alguna hipótesis; otro ejemplo: el cálculo no sólo sigue a la observación sino que siempre es indispensable para planearla y registrarla. La racionalidad y objetividad del conocimiento científico pueden analizarse en un cúmulo de características a las que pasaremos revista en lo que sigue.

3. Inventario de las principales características de la ciencia fáctica

1) *El conocimiento científico es fáctico*: parte de los hechos, los respeta hasta cierto punto y siempre vuelve a ellos. La ciencia intenta descri-

bir los hechos tales como son, independientemente de su valor emocional o comercial: la ciencia no poetiza los hechos ni los vende, si bien sus hazañas son una fuente de poesía y de negocios. En todos los campos, la ciencia comienza estableciendo los hechos: esto requiere curiosidad impersonal, desconfianza por la opinión prevaleciente y sensibilidad a la novedad.

Los enunciados fácticos confirmados se llaman usualmente "datos empíricos": se obtienen con ayuda de teorías (por esquemáticas que sean) y son a su vez la materia prima de la elaboración teórica. Una subclase de datos empíricos es de tipo cuantitativo; los datos numéricos y métricos se disponen a menudo en tablas, las más importantes de las cuales son las tablas de constantes (tales como las de los puntos de fusión de las diferentes sustancias). Pero la recolección de datos y su ulterior disposición en tablas no es la finalidad principal de la investigación: la información de esta clase debe incorporarse a teorías si ha de convertirse en una herramienta para la inteligencia y la aplicación. ¿De qué sirve conocer el peso específico del hierro si carecemos de fórmulas mediante las cuales podemos relacionarlo con otras cantidades?

No siempre es posible, ni siquiera deseable, respetar enteramente los hechos cuando se los analiza, y no hay ciencia sin análisis, aun cuando el análisis no sea sino un medio para la reconstrucción final de los todos. El físico atómico perturba el átomo al que desea espiar; el biólogo mo-

difica y puede incluso matar al ser vivo que analiza; el antropólogo empeñado en el estudio de campo de una comunidad provoca en ella ciertas modificaciones. Ninguno de ellos aprehende su objeto tal como es, sino tal como queda modificado por sus propias operaciones; sin embargo, en todos los casos tales cambios son objetivos, y se presume que pueden entenderse en términos de leyes: no son conjurados arbitrariamente por el experimentador. Más aún, en todos los casos el investigador intenta describir las características y el monto de la perturbación que produce en el acto del experimento; procura, en suma, estimar la desviación o "error" producido por su intervención activa. Porque los científicos actúan haciendo tácitamente la suposición de que el mundo existiría aun en su ausencia, aunque, desde luego, no exactamente de la misma manera.

2) *El conocimiento científico trasciende los hechos:* descarta hechos, produce nuevos hechos y los explica. El sentido común parte de los hechos y se atiene a ellos: a menudo se limita al hecho aislado, sin ir muy lejos en el trabajo de correlacionarlo con otros o de explicarlo. En cambio, la investigación científica no se limita a los hechos observados: los científicos exprimen la realidad a fin de ir más allá de las apariencias; rechazan el grueso de los hechos percibidos, por ser un montón de accidentes; seleccionan los que consideran que son relevantes; controlan hechos y, en lo posible, los reproducen. Incluso producen cosas nuevas, desde instrumentos hasta partículas elemen-

tales; obtienen nuevos compuestos químicos, nuevas variedades vegetales y animales y, al menos en principio, crean nuevas pautas de conducta individual y social.

Más aún, los científicos usualmente no aceptan nuevos hechos a menos que puedan certificar de alguna manera su autenticidad; y esto se hace, no tanto contrastándolos con otros hechos, cuanto mostrando que son compatibles con lo que se sabe. Los científicos descartan las imposturas y los trucos mágicos porque no encuadran en hipótesis muy generales y fidedignas, que han sido puestas a prueba en incontables ocasiones. Vale decir, los científicos no consideran su propia experiencia individual como un tribunal inapelable; se fundan, en cambio, en la experiencia colectiva y en la teoría.

Hay más: el conocimiento científico racionaliza la experiencia en lugar de limitarse a describirla; la ciencia da cuenta de los hechos, no inventariándolos sino explicándolos por medio de hipotesis (en particular, enunciados de leyes) y sistemas de hipótesis (teorías). Los científicos conjeturan lo que hay tras los hechos observados, y de continuo inventan conceptos (tales como los de átomo, campo, masa, energía, adaptación, integración, selección, clase social o tendencia histórica) que carecen de correlato empírico, esto es, que no corresponden a perceptos, aun cuando presumiblemente se refieren a cosas, cualidades o relaciones existentes objetivamente. No percibimos los campos eléctricos o las clases sociales: conjeturamos

su existencia a partir de hechos experimentables, y tales conceptos son significativos tan sólo en ciertos contextos teóricos.

Este trascender la experiencia inmediata, este salto del nivel observacional al teórico, le permite a la ciencia mirar con desconfianza los enunciados sugeridos por meras coincidencias; le permite predecir la existencia real de cosas y procesos ocultos a primera vista, pero que instrumentos (materiales o conceptuales) más potentes pueden descubrir. Las discrepancias entre las previsiones teóricas y los hallazgos empíricos figuran entre los estímulos más fuertes para edificar teorías nuevas y diseñar nuevos experimentos. No son los hechos por sí mismos sino su elaboración teórica y la comparación de las consecuencias de las teorías con los datos observacionales, la principial fuente del descubrimiento de nuevos hechos.

3) *La ciencia es analítica*: la investigación científica aborda problemas circunscriptos, uno a uno, y trata de descomponerlo todo en elementos (no necesariamente últimos o siquiera reales). La investigación científica no se plantea cuestiones tales como "¿Cómo es el universo en su conjunto?", o "¿Cómo es posible el conocimiento?" Trata, en cambio, de entender toda situación total en términos de sus componentes; intenta descubrir los elementos que componen cada totalidad, y las interconexiones que explican su integración.

Los problemas de la ciencia son parciales y así son también, por consiguiente, sus soluciones; pero, más aún: al comienzo los problemas son es-

trechos o es preciso estrecharlos. Pero, a medida que la investigación avanza, su alcance se amplía. Los resultados de la ciencia son generales, tanto en el sentido de que se refieren a clases de objetos (p. ej.: lluvia), como en el de que están o tienden a ser incorporados en síntesis conceptuales llamadas teorías. El análisis, tanto de los problemas como de las cosas, no es tanto un objetivo como una herramienta para construir síntesis teóricas. La ciencia auténtica no es atomista ni totalista, sino sistémica.

La investigación comienza descomponiendo sus objetos a fin de descubrir el "mecanismo" interno responsable de los fenómenos observados. Pero el desmontaje del "mecanismo" no se detiene cuando se ha investigado la naturaleza de sus partes; el próximo paso es el examen de la interdependencia de las partes, y la etapa final es la tentativa de reconstruir el todo en términos de sus partes interconectadas. El análisis no acarrea el descuido de la totalidad; lejos de disolver la integración, el análisis es la única manera conocida de descubrir cómo emergen, subsisten y se desintegran los todos. La ciencia no ignora la síntesis; lo que sí rechaza es la pretensión irracionalista de que las síntesis pueden ser aprehendidas por una intuición especial, sin previo análisis.

4) *La investigación científica es especializada*: una consecuencia del enfoque analítico de los problemas es la especialización. No obstante la unidad del método científico, su aplicación depende, en gran medida, del asunto; esto explica la

multiplicidad de técnicas y la relativa independencia de los diversos sectores de la ciencia.

Sin embargo, es menester no exagerar la diversidad de las ciencias al punto de borrar su unidad metodológica. El viejo dualismo materia/espíritu había sugerido la división de las ciencias en *Naturwissenschaften*, o ciencias de la naturaleza, y *Geisteswissenschaften*, o ciencias del espíritu. Pero estos géneros difieren en cuanto al asunto, a las técnicas y al grado de desarrollo, no así en lo que respecta al objetivo, método y alcance. Además, la división en cuestión es inválida porque hay ciencias híbridas, es decir, socionaturales, tales como la psicología, la antropología, la lingüística y la demografía.

El dualismo razón/experiencia había sugerido, a su vez, la división de las ciencias fácticas en racionales y empíricas, que tampoco es adecuada porque todas son a la vez racionales y empíricas. Menos sostenible aún es la dicotomía ciencias deductivas/ciencias inductivas, ya que toda empresa científica —sin excluir el dominio de las ciencias formales— es tan inductiva como deductiva, sin hablar de otros tipos de inferencia, por ejemplo la analógica.

La especialización no ha impedido la formación de campos interdisciplinarios, tales como la biofísica, la bioquímica, la psicofisiología, la psicología social, la teoría de la información, la cibernética o la investigación operacional. Con todo, la especialización tiende a estrechar la visión del científico individual. Un único remedio ha resulta-

do eficaz contra la unilateralidad profesional, y es una dosis de filosofía.

5) *El conocimiento científico es claro y preciso:* sus problemas son distintos, sus resultados son claros. El conocimiento ordinario, en cambio, usualmente es vago e inexacto; en la vida diaria nos preocupamos poco por dar definiciones precisas, descripciones exactas o mediciones afinadas: si éstas nos preocuparan demasiado, no lograríamos marchar al paso de la vida. La ciencia torna preciso lo que el sentido común conoce de manera nebulosa; pero, desde luego, la ciencia es mucho más que sentido común organizado: aunque proviene del sentido común, la ciencia constituye una rebelión contra su vaguedad y superficialidad. El conocimiento científico procura la precisión; nunca está enteramente libre de vaguedades, pero se las ingenia para mejorar la exactitud; nunca está del todo libre de error, pero posee una técnica única para encontrar errores y para sacar provecho de ellos.

La claridad y la precisión en ciencia se obtienen de las siguientes maneras:

a) los problemas se formulan de manera clara; lo primero, y a menudo lo más difícil, es distinguir cuáles son los problemas; ni hay artillería analítica o experimental que pueda ser eficaz si no se ubica adecuadamente al enemigo;

b) la ciencia parte de nociones que parecen claras al no iniciado; y las complica, purifica y eventualmente las rechaza; la transformación progresiva de las nociones corrientes se efectúa incluyéndolas en esquemas teóricos. Así, por ejemplo,

"distancia" adquiere un sentido preciso al ser incluida en la geometría métrica y en la física;

c) la ciencia define la mayoría de sus conceptos: algunos de ellos se definen en términos de conceptos no definidos o primitivos; otros, de manera implícita, esto es, por la función que desempeñan en un sistema teórico (definición contextual). Las definiciones son convencionales, pero no se las elige caprichosamente: deben ser convenientes y fértiles. (¿De qué vale, por ejemplo, poner un nombre especial a las muchachas pecosas que estudian ingeniería y pesan más de 50 kg?) Una vez que se ha elegido una definición, el discurso restante debe guardarle fidelidad si se quiere evitar inconsecuencias;

d) la ciencia crea lenguajes artificiales inventando símbolos (palabras, signos matemáticos, símbolos químicos, etc.); a estos signos se les atribuye significados determinados por medio de reglas de designación (tal como "en el presente contexto «H» designa el elemento de peso atómico unitario"). Los símbolos básicos serán tan simples como sea posible, pero podrán combinarse conforme a reglas determinadas para formar configuraciones tan complejas como sea necesario (las leyes de combinación de los signos que intervienen en la producción de expresiones complejas se llaman *reglas de formación*);

e) la ciencia procura siempre medir y registrar los fenómenos. Los números y las formas geométricas son de gran importancia en el registro, la descripción y la inteligencia de los sucesos

y procesos. En lo posible, tales datos debieran disponerse en tablas o resumirse en fórmulas matemáticas. Sin embargo, la formulación matemática, deseable como es, no es una condición indispensable para que el conocimiento sea científico; lo que caracteriza el conocimiento científico es la exactitud en un sentido general antes que la exactitud numérica o métrica, la que es inútil si media la vaguedad conceptual o la inadecuación a la realidad. Más aún, la investigación científica emplea, en medida creciente, capítulos no numéricos y no métricos de la matemática, tales como la topología, la teoría de los grupos o el álgebra de las clases, que no son ciencias del número y de la figura, sino de la relación.

6) *El conocimiento científico es comunicable*: no es inefable sino expresable, no es privado sino público. El lenguaje científico comunica información a quienquiera haya sido adiestrado para entenderlo. Hay, ciertamente, sentimientos oscuros y nociones difusas, incluso en el desarrollo de la ciencia (aunque no en la presentación final del trabajo científico); pero es preciso aclararlos antes de poder estimar su adecuación. Lo que es inefable puede ser propio de la poesía o de la música, no de la ciencia, cuyo lenguaje es informativo y no expresivo o imperativo. La inefabilidad misma es, en cambio, tema de investigación científica, sea psicológica o lingüística.

La comunicabilidad es posible gracias a la precisión; y es a su vez una condición necesaria para la verificación de los datos empíricos y de las

hipótesis científicas. Aun cuando, por motivos comerciales o políticos, se mantengan en secreto durante un tiempo algunos trozos del saber, deben ser comunicables en principio para que puedan ser considerados científicos. La comunicación de los resultados y de las técnicas de la ciencia no sólo perfecciona la educación general sino que multiplica las posibilidades de su confirmación o refutación. La verificación independiente ofrece las máximas garantías técnicas y morales, y ahora es posible, en muchos campos, en escala internacional. Por esto, los científicos consideran el secreto en materia científica como enemigo del progreso de la ciencia; la política del secreto científico es, en efecto, el más eficaz originador del estancamiento en la cultura, en la tecnología y en la economía, así como una fuente de corrupción moral.

7) *El conocimiento científico es verificable:* debe aprobar el examen de la experiencia. A fin de explicar un conjunto de fenómenos, el científico inventa conjeturas fundadas de alguna manera en el saber adquirido. Sus suposiciones pueden ser cautas o audaces, simples o complejas; en todo caso, deben ser puestas a prueba. El *test* de las hipótesis fácticas es empírico, esto es, observacional o experimental. El haberse dado cuenta de esta verdad hoy tan trillada es la contribución inmortal de la ciencia helenística. En este sentido, las ideas científicas (incluidos los enunciados de leyes) no son superiores a las herramientas o a los vestidos: si fracasan en la práctica, fracasan por entero.

La experimentación puede calar más profundamente que la observación, porque efectúa cambios en lugar de limitarse a registrar variaciones: aísla y controla las variables sensibles o pertinentes. Sin embargo, los resultados experimentales son pocas veces interpretables de una sola manera. Más aún, no todas las ciencias pueden experimentar; y en ciertos capítulos de la astronomía y de la economía se alcanza una gran exactitud sin ayuda del experimento. La ciencia fáctica es por esto *empírica* en el sentido de que la comprobación de sus hipótesis involucra la experiencia; pero no es necesariamente *experimental* y, en particular, no es agotada por las ciencias de laboratorio, tales como la física.

La prescripción de que las hipótesis científicas deben ser capaces de aprobar el examen de la experiencia es una de las reglas del método científico; la aplicación de esta regla depende del tipo de objeto, del tipo de la hipótesis en cuestión y de los medios disponibles. Por esto, se necesita una multitud de técnicas de verificación empírica. La verificación de la fórmula de un compuesto químico se hace de una manera muy diferente de la verificación de un cálculo astronómico o de una hipótesis concerniente al pasado de las rocas o de los hombres. Las técnicas de verificación evolucionan en el curso del tiempo; sin embargo, siempre consisten en poner a prueba consecuencias particulares de hipótesis generales (entre ellas, enunciados de leyes). Siempre se reducen a mostrar que hay, o que no hay, algún fundamento para creer que las

suposiciones en cuestión corresponden a los hechos observados o a los valores medidos.

La verificabilidad hace a la esencia del conocimiento científico; si así no fuera, no podría decirse que los científicos procuran alcanzar conocimiento objetivo.

8) *La investigación científica es metódica*: no es errática sino planeada. Los investigadores no tantean en la oscuridad: saben lo que buscan y cómo encontrarlo. El planeamiento de la investigación no excluye el azar; sólo que, al hacer un lugar a los acontecimientos imprevistos, es posible aprovechar la interferencia del azar y la novedad inesperada. Más aún, a veces el investigador produce el azar deliberadamente. Por ejemplo, para asegurar la uniformidad de una muestra, y para impedir una preferencia inconsciente en la elección de sus miembros, a menudo se emplea la técnica de la casualización, en que la decisión acerca de los individuos que han de formar parte de ciertos grupos se deja librada a una moneda o a algún otro dispositivo. De esta manera, el investigador pone el azar al servicio del orden; en lo cual no hay paradoja, porque el acaso opera al nivel de los individuos, al par que el orden opera en el grupo como totalidad.

Todo trabajo de investigación se funda sobre el conocimiento anterior, y en particular sobre las conjeturas mejor confirmadas. (Uno de los muchos problemas de la metodología es, precisamente, averiguar cuáles son los criterios para decidir si una hipótesis dada puede considerarse razonable-

mente confirmada, esto es, si el peso que le acuerdan los fundamentos inductivos y de otro orden basta para conservarla). Más aún, la investigación procede conforme a reglas y técnicas que han resultado eficaces en el pasado, pero que son perfeccionadas continuamente, no sólo a la luz de nuevas experiencias, sino también de resultas del examen matemático y filosófico. Una de las reglas de procedimiento de la ciencia fáctica es la siguiente: las variables pertinentes (o que se sospecha que son sensibles) debieran variarse una por vez.

La ciencia fáctica emplea el método experimental concebido en un sentido amplio. Este método consiste en el *test* empírico de conclusiones particulares extraídas de hipótesis generales (tales como "Los gases se dilatan cuando se los calienta" o "Los hombres se rebelan cuando se los oprime"). Este tipo de verificación requiere la manipulación, la observación y el registro de fenómenos; requiere también el control de las variables o factores pertinentes; siempre que sea posible debiera incluir la producción artificial deliberada de los fenómenos en cuestión, y en todos los casos exige el análisis de los datos obtenidos en el curso de los procedimientos empíricos. Los datos aislados y crudos son inútiles y no son dignos de confianza; es preciso elaborarlos, organizarlos y confrontarlos con las conclusiones teóricas.

El método científico no provee recetas infalibles para encontrar la verdad: sólo contiene un conjunto de prescripciones falibles (y perfectibles) para el planeamiento de observaciones y experi-

mentos, para la interpretación de sus resultados, y para el planteo mismo de los problemas. Es, en suma, la manera en que la ciencia explora lo desconocido. Subordinadas a las reglas generales del método científico, y al mismo tiempo en apoyo de ellas, encontramos las diversas técnicas que se emplean en las ciencias especiales: las técnicas para pesar, para observar por el microscopio, para analizar compuestos químicos, para dibujar gráficos que resumen datos empíricos, para reunir informaciones acerca de costumbres, etc. La ciencia es, pues, esclava de sus propios métodos y técnicas mientras éstos tienen éxito; pero es libre de multiplicar y de modificar en todo momento sus reglas, en aras de una mayor racionalidad y objetividad.

9) *El conocimiento científico es sistemático*: una ciencia no es un agregado de informaciones inconexas, sino un sistema de ideas conectadas lógicamente entre sí. Todo sistema de ideas, caracterizado por cierto conjunto básico (pero refutable) de hipótesis peculiares, y que procura adecuarse a una clase de hechos, es una teoría. Todo capítulo de una ciencia especial contiene teorías o sistemas de ideas que están relacionadas lógicamente entre sí, esto es, que están ordenadas mediante la relación "implica". (Ejemplo: "Si aumenta la desocupación aumenta el delito".) Esta conexión entre las ideas puede calificarse de orgánica, en el sentido de que la sustitución de cualquiera de las hipótesis básicas produce un cambio radical en la teoría o grupo de teorías.

El fundamento de una teoría dada no es un conjunto de hechos, sino, más bien, un conjunto de principios o hipótesis de cierto grado de generalidad (y, por consiguiente, de cierta fertilidad lógica). Las conclusiones (o teoremas) pueden extraerse de los principios, sea en la forma natural o con ayuda de técnicas especiales que involucran operaciones matemáticas.

El carácter matemático del conocimiento científico —esto es, el hecho de que es fundado, ordenado y coherente— es lo que lo hace racional. La racionalidad permite que el progreso científico se efectúe no sólo por la acumulación gradual de resultados, sino también por revoluciones. Las revoluciones científicas no son descubrimientos de nuevos hechos aislados, ni son perfeccionamientos en la exactitud de las observaciones, sino que consisten en la sustitución de hipótesis de gran alcance (principios) por nuevos axiomas, y en el reemplazo de teorías enteras por otros sistemas teóricos. Sin embargo, semejantes revoluciones son a menudo provocadas por el descubrimiento de nuevos hechos de los que no dan cuenta las teorías anteriores, aunque a veces se encuentran en el proceso de comprobación de dichas teorías; y las nuevas teorías se tornan verificables en muchos casos, merced a la invención de nuevas técnicas de medición de mayor precisión.

10) *El conocimiento científico es general*: ubica los hechos singulares en pautas generales, los enunciados particulares en esquemas amplios. El científico se ocupa del hecho singular en la medi-

da en que éste es miembro de una clase o caso de una ley; más aún, presupone que todo hecho es clasificable y legal. No es que la ciencia ignore la cosa individual o el hecho irrepetible; lo que ignora es el hecho aislado. Por esto la ciencia no se sirve de los datos empíricos —que siempre son singulares— como tales; éstos son mudos mientras no se los manipula y convierte en piezas de sistemas teóricos.

En efecto, uno de los principios ontológicos que subyacen a la investigación científica es que la variedad y aun la unicidad en algunos respectos son compatibles con la uniformidad y la generalidad en otros respectos. Al químico no le interesa ésta o aquella hoguera, sino el proceso de la combustión en general; trata de descubrir lo que comparten todos los singulares. El científico intenta exponer los universales que se esconden en el seno de los propios singulares; es decir, no considera los universales *ante rem* ni *post rem* sino *in re*, en la cosa, y no antes o después de ella. Los escolásticos medievales clasificarían al científico moderno como realista inmanentista porque, al descartar los detalles, al procurar descubrir los rasgos comunes a individuos que son únicos en otros respectos, al buscar las variables pertinentes (en particular las cualidades esenciales) y las relaciones constantes entre ellas (las leyes), el científico intenta exponer la naturaleza esencial de las cosas naturales y humanas.

El lenguaje científico no contiene solamente términos que designan hechos singulares y expe-

riencias individuales, sino también términos generales que se refieren a clases de hechos. La generalidad del lenguaje de la ciencia no tiene, sin embargo, el propósito de alejar a la ciencia de la realidad concreta: por el contrario, la generalización es el único medio que se conoce para adentrarse en lo concreto, para apresar la esencia de las cosas (sus cualidades y leyes esenciales). Con esto, el científico evita en cierta medida las confusiones y los engaños provocados por el flujo deslumbrador de los fenómenos. Tampoco se asfixia la utilidad en la generalidad: por el contrario, los esquemas generales de la ciencia encuadran una cantidad ilimitada de casos específicos, proveen leyes de amplio alcance que incluyen y corrigen todas las recetas válidas del sentido común y de la técnica precientífica.

11) *El conocimiento científico es legal*: busca leyes (de la naturaleza y de la sociedad) y las aplica. El conocimiento científico inserta los hechos singulares en pautas generales llamadas "leyes naturales" o "leyes sociales". Tras el desorden y la fluidez de las apariencias, la ciencia fáctica descubre las pautas regulares de la estructura y del proceso del ser y del devenir. En la medida en que la ciencia es legal, es esencialista: intenta llegar a la raíz de las cosas. Encuentra la esencia en variables pertinentes y en las relaciones invariantes entre ellas.

Hay leyes de hechos y leyes mediante las cuales se pueden explicar otras leyes. El principio de Arquímedes pertenece a la primera clase; pero a su vez puede deducirse de los principios generales

de la mecánica; por consiguiente, ha dejado de ser un principio independiente, y ahora es un teorema deducible de hipótesis de nivel más elevado. Las leyes de la física proveen la base de las leyes de las combinaciones químicas; las leyes de la fisiología explican ciertos fenómenos psíquicos; y las leyes de la psicología y de la sociología pertenecen a los fundamentos de la economía. Es decir, los enunciados de las leyes se organizan en una estructura de niveles.

Ciertamente, los enunciados de las leyes son transitorios; pero ¿son inmutables las leyes mismas? Si se consideran las leyes como las pautas mismas del ser y del devenir, entonces debieran cambiar junto con las cosas mismas; por lo menos, debe admitirse que, al emerger nuevos niveles, sus cualidades peculiares se relacionan entre sí mediante nuevas leyes. Por ejemplo, las leyes de la economía han emergido en el curso de la historia sobre la base de otras leyes (biológicas y psicológicas) y, más aún, algunas de ellas cambian con el tipo de organización social.

Por supuesto, no todos los hechos singulares conocidos han sido ya convertidos en casos particulares de leyes generales; en particular, los sucesos y procesos de los niveles superiores han sido legalizados sólo en pequeña medida. Pero esto se debe en parte al antiguo prejuicio de que lo humano no es legal, así como a la antigua creencia pitagórica de que solamente las relaciones numéricas merecen llamarse "leyes científicas". Debiera emplearse el *stock* íntegro de las herramientas con-

ceptuales en la búsqueda de las leyes de la mente y de la cultura; más aún, acaso el *stock* de que se dispone es insuficiente y sea preciso inventar herramientas radicalmente nuevas para tratar los fenómenos mentales y culturales, tal como el nacimiento de la mecánica moderna hubiera sido imposible sin la invención expresa del cálculo infinitesimal.

Pero el ulterior avance en el progreso de la legalización de los fenómenos no físicos requiere, por sobre todo, una nueva actitud frente al concepto mismo de ley científica. En primer lugar, es preciso comprender que hay muchos tipos de leyes (aun dentro de una misma ciencia), ninguno de los cuales es necesariamente mejor que los tipos restantes. En segundo lugar, debiera tornarse un lugar común entre los científicos de la cultura el que las leyes no se encuentran por mera observación y el simple registro, sino poniendo a prueba hipótesis: los enunciados de leyes no son, en efecto, sino hipótesis confirmadas. Y ¿cómo habríamos de emprender la confección de hipótesis científicas si no presumiéramos que todo hecho singular es legal?

12) *La ciencia es explicativa:* intenta explicar los hechos en términos de leyes, y las leyes en términos de principios. Los científicos no se conforman con descripciones detalladas; además de inquirir cómo son las cosas, procuran responder a porqués: por qué ocurren los hechos como ocurren y no de otra manera. La ciencia deduce proposiciones relativas a hechos singulares a partir de leyes generales, y deduce las leyes a partir de enuncia-

dos nomológicos aun más generales (principios). Por ejemplo, las leyes de Kepler explicaban una colección de hechos observados del movimiento planetario; y Newton explicó esas leyes deduciéndolas de principios generales, explicación que permitió a otros astrónomos dar cuenta de las irregularidades de las órbitas de los planetas, que eran desconocidas para Kepler.

Solía creerse que explicar es señalar la causa, pero en la actualidad se reconoce que la explicación causal no es sino un tipo de explicación científica. La explicación científica se efectúa siempre en términos de leyes, y las leyes causales no son sino una subclase de las leyes científicas. Hay diversos tipos de leyes científicas y, por consiguiente, hay una variedad de tipos de explicaciones científicas: morfológicas, cinemáticas, dinámicas, de composición, de conservación, de asociación, etcétera.

La historia de la ciencia enseña que las explicaciones científicas se corrigen o descartan sin cesar. ¿Significa esto que son todas falsas? En la ciencia fáctica, la verdad y el error no son del todo ajenos entre sí: hay verdades parciales y errores parciales; hay aproximaciones buenas y otras malas. La ciencia no obra como Penélope, sino que emplea la tela tejida ayer. Las explicaciones científicas no son finales pero son perfectibles.

13) *El conocimiento científico es predictivo*: trasciende la masa de los hechos de experiencia, imaginando cómo puede haber sido el pasado y cómo podrá ser el futuro. La predicción es, en pri-

mer lugar, una manera eficaz de poner a prueba las hipótesis; pero también es la clave del control o aun de la modificación del curso de los acontecimientos. La predicción científica, en contraste con la profecía, se funda sobre leyes y sobre informaciones específicas fidedignas, relativas al estado de cosas actual o pasado. No es del tipo "Ocurrirá e", sino más bien de este otro: "Ocurrirá e (de la clase E) siempre que suceda c (de la clase C), pues todo hecho de clase C es acompañado o seguido de un hecho de clase E".

La predicción científica se caracteriza por su perfectibilidad antes que por su certeza. Más aún, las predicciones que se hacen con la ayuda de reglas empíricas son a veces más exactas que las predicciones penosamente elaboradas con herramientas científicas (leyes, informaciones específicas y deducciones); tal es el caso, con frecuencia, de los pronósticos meteorológicos, de la prognosis médica y de la profecía política. Pero, en tanto que la profecía no es perfectible y no puede usarse para poner a prueba hipótesis, la predicción es perfectible y, si falla, nos obliga a corregir nuestras suposiciones, alcanzando así una inteligencia más profunda. Por esto la profecía exitosa no es un aporte al conocimiento teórico, en tanto que la predicción científica fallida puede contribuir a él.

Puesto que la predicción científica depende de leyes y de ítems de información específica, puede fracasar por inexactitud de los enunciados de las leyes o por imprecisión de la información disponi-

ble. (También puede fallar, por supuesto, debido a errores cometidos en el proceso de inferencia lógica o matemática que conduce de las premisas [leyes e informaciones] a la conclusión [enunciado predictivo]). Una fuente importante de fallas en la predicción es el conjunto de suposiciones acerca de la naturaleza del objeto (sistema físico, organismo vivo, grupo social, etc.) cuyo comportamiento ha de predecirse. Por ejemplo, puede ocurrir que creamos que el sistema en cuestión está suficientemente aislado de las perturbaciones exteriores, cuando en rigor éstas cuentan a la larga; dado que la aislación es una condición necesaria de la descripción del sistema con ayuda de un puñado de enunciados de leyes, no debiera sorprender que sea tan difícil predecir el comportamiento de sistemas abiertos tales como el océano, la atmósfera, el ser vivo o la economía.

Puesto que la predicción científica se funda en leyes científicas, hay tantas clases de predicciones como clases de enunciados nomológicos. Algunas leyes nos permiten predecir resultados individuales, aunque no sin error si la predicción se refiere al valor de una cantidad. Otras leyes, incapaces de decirnos nada acerca del comportamiento de los individuos (átomos, personas, etc.) son en cambio la base para la predicción de algunas tendencias globales y propiedades colectivas de colecciones numerosas de elementos similares: son las leyes estadísticas. Las leyes sociales son de este tipo; por esto es casi imposible la predicción de los sucesos individuales en el campo de la historia,

pudiendo preverse solamente el curso general de los acontecimientos.

14) *La ciencia es abierta*: no reconoce barreras *a priori* que limiten el conocimiento. Si un conocimiento fáctico no es refutable en principio, entonces no pertenece a la ciencia sino a algún otro campo. Las nociones acerca de nuestro medio natural o social, o acerca del yo, no son finales: están todas en movimiento, todas son falibles. Siempre es concebible que pueda surgir una nueva situación (nuevas informaciones o nuevos trabajos teóricos) en que nuestras ideas, por firmemente establecidas que parezcan, resulten inadecuadas en algún sentido. La ciencia carece de axiomas evidentes; incluso los principios más generales y seguros son postulados que pueden ser corregidos o reemplazados. A consecuencia del carácter hipotético de los enunciados de leyes y de la naturaleza perfectible de los datos empíricos, la ciencia no es un sistema dogmático y cerrado sino controvertido y abierto. O, más bien, la ciencia es abierta como sistema porque es falible y por consiguiente capaz de progresar. En cambio, puede argüirse que la ciencia es metodológicamente cerrada, no en el sentido de que las reglas del método científico sean finales, sino en el sentido de que es autocorrectiva: el requisito de la verificabilidad de las hipótesis científicas basta para asegurar el progreso científico.

Tan pronto como ha sido establecida una teoría científica, corre el peligro de ser refutada o, al menos, de que se circunscriba su dominio. Un sis-

tema cerrado de conocimiento fáctico, que excluya toda ulterior investigación, puede llamarse sabiduría pero es en rigor un detritus de la ciencia. El sabio moderno, a diferencia del antiguo, no es tanto un acumulador de conocimientos como un generador de problemas. Por consiguiente, prefiere los últimos números de las revistas especializadas a los manuales, aun cuando estos últimos sean depósitos de verdad más vastos y fidedignos que aquéllas. El investigador moderno ama la verdad pero no se interesa por las teorías irrefutables. Una teoría puede haber permanecido intocada no tanto por su alto contenido de verdad cuanto porque nadie la ha usado. No se necesita emprender una investigación empírica para probar la tautología de que ni siquiera los científicos están casados con solteronas.

Los modernos sistemas de conocimiento científico son como organismos en crecimiento: mientras están vivos cambian sin pausa. Ésta es una de las razones por las cuales la ciencia es éticamente valiosa: porque nos recuerda que la corrección de errores es tan valiosa como el no cometerlos, y que probar cosas nuevas e inciertas es preferible a rendir culto a las viejas y garantidas. La ciencia, como los organismos, cambia a la vez internamente y debido a sus contactos con sus vecinos; esto es, resolviendo sus problemas específicos y siendo útil en otros campos.

15) *La ciencia es útil*: porque busca la verdad, la ciencia es eficaz en la provisión de herramientas para el bien y para el mal. El conocimiento

ordinario se ocupa usualmente de lograr resultados capaces de ser aplicados en forma inmediata; con ello no es suficientemente verdadero, con lo cual no puede ser suficientemente eficaz. Cuando se dispone de un conocimiento adecuado de las cosas es posible manipularlas con éxito. La utilidad de la ciencia es una consecuencia de su objetividad: sin proponerse necesariamente alcanzar resultados aplicables, la investigación los provee a la corta o a la larga. La sociedad moderna paga la investigación porque ha aprendido que la investigación rinde. Por este motivo, es redundante exhortar a los científicos a que produzcan conocimientos aplicables: no pueden dejar de hacerlo. Es cosa de los técnicos emplear el conocimiento científico con fines prácticos, y los políticos y empresarios son los responsables de que la ciencia y la tecnología se empleen en beneficio de la humanidad. Los científicos pueden a lo sumo, aconsejar acerca de cómo puede hacerse uso racional, eficaz y bueno de la ciencia.

La técnica precientífica era primordialmente una colección de recetas pragmáticas no entendidas, muchas de las cuales desempeñaban la función de ritos mágicos. La técnica moderna es, en medida creciente —aunque no exclusivamente—, ciencia aplicada. La ingeniería es física y química aplicadas, la medicina es biología aplicada, la psiquiatría es psicología y neurología aplicadas; y debiera llegar el día en que la política se convierta en sociología y economía aplicadas.

Pero la tecnología es más que ciencia aplicada: en primer lugar, porque tiene sus propios procedimientos de investigación, adaptados a circunstancias concretas que distan de los casos puros que estudia la ciencia. En segundo lugar, porque toda rama de la tecnología contiene un cúmulo de reglas empíricas descubiertas antes que los principios científicos en los que —si dichas reglas se confirman— terminan por ser absorbidas. La tecnología no es meramente el resultado de aplicar el conocimiento científico existente a los problemas prácticos: la tecnología viva es, esencialmente, el enfoque científico de problemas prácticos, es decir, el tratamiento de estos problemas sobre un fondo de conocimiento científico y con ayuda del método científico. Por esto la tecnología, sea de las cosas o de los hombres, es fuente de conocimientos nuevos.

La conexión de la ciencia con la tecnología no es, por consiguiente, asimétrica. Todo avance tecnológico plantea problemas científicos cuya solución puede consistir en la invención de nuevas teorías o de nuevas técnicas de investigación que conduzcan a un conocimiento más adecuado y a un mejor dominio del asunto. La ciencia y la tecnología constituyen un ciclo de sistemas interactuantes que se alimentan el uno al otro. El científico torna inteligible lo que hace el técnico y éste provee a la ciencia de instrumentos y de comprobaciones; y, lo que es igualmente importante, el técnico no cesa de formular preguntas al científico, añadiendo así un motor externo al motor inter-

no del progreso científico. La continuación de la vida sobre la Tierra depende del ciclo del carbono: los animales se alimentan de plantas, las que a su vez obtienen su carbono de lo que exhalan los animales. Análogamente, la continuación de la civilización moderna depende, en gran medida, del ciclo del conocimiento: la tecnología moderna come ciencia, y la ciencia moderna depende a su vez del equipo y del estímulo que le provee una industria altamente tecnificada.

Pero la ciencia es útil en más de una manera. Además de constituir el fundamento de la tecnología, la ciencia es útil en la medida en que se la emplea en la edificación de concepciones del mundo que concuerdan con los hechos, y en la medida en que crea el hábito de adoptar una actitud de libre y valiente examen, en que acostumbra a la gente a poner a prueba sus afirmaciones y a argumentar correctamente. No menor es la utilidad que presta la ciencia como fuente de apasionantes rompecabezas filosóficos, y como modelo de la investigación filosófica.

En resumen, la ciencia es valiosa como herramienta para domar la naturaleza y remodelar a la sociedad; es valiosa en sí misma, como clave para la inteligencia del mundo y del yo; y es eficaz en el enriquecimiento, la disciplina y la liberación de nuestra mente.

¿Cuál es el método de la ciencia?

The lame in the path outstrips the swift
who wander from it.

F. Bacon.

1. La ciencia, conocimiento verificable

En su deliciosa biografía del Dante (c. 1360), Boccaccio[1] expuso su opinión —que no viene al caso— acerca del origen de la palabra "poesía", concluyendo con este comentario: "otros lo atribuyen a razones diferentes, acaso aceptables; pero ésta me *gusta más*". El novelista aplicaba, al *conocimiento* acerca de la poesía y de su nombre, el mismo criterio que podría usarse para apreciar la poesía misma: el gusto. Confundía así valores situados en niveles diferentes: el estético, perteneciente a la esfera de la sensibilidad, y el gnoseológico, que no obstante estar enraizado en la sensibilidad está enriquecido con una cualidad emergente: la razón.

Semejante confusión no es exclusiva de poetas; incluso Hume, en una obra célebre por su crítica mortífera de varios dogmas tradicionales, escogió el gusto como criterio de verdad. En su *Treatise of Human Nature*[2] (1739) puede leerse: "¡No es sólo en poesía y en música que *debemos se-*

---

[1] G. Boccaccio, *Vita di Dante*, en *Il comento alla Divina Commedia e gli altri scritti intorno a Dante* (Bari, Laterza, 1918), I, p. 37. Subrayado mío.

[2] D. Hume, *A Treatise of Human Nature* (London, Everyman, 1911), I, p. 105. Subrayado mío.

guir nuestro gusto, sino también en la filosofía [que en aquella época incluía habitualmente la ciencia]. Cuando estoy convencido de algún principio, no es sino una idea que me golpea [*strikes*] con mayor fuerza. Cuando prefiero un conjunto de argumentos por sobre otros, no hago sino decidir, sobre la base de mi sentimiento, acerca de la superioridad de su influencia". El subjetivismo era así la playa en que desembarcaba la teoría psicologista de las "ideas" inaugurada por el empirismo de Locke.

El recurso al gusto no era, por supuesto, peor que el argumento de autoridad, criterio de verdad que ha mantenido enjaulado el pensamiento durante tanto tiempo y con tanta eficacia. Desgraciadamente, la mayoría de la gente, y hasta la mayoría de los filósofos, aún creen —u obran como si creyeran— que la manera correcta de decidir el valor de verdad de un enunciado es someterlo a la prueba de algún texto: es decir, verificar si es compatible con (o deducible de) frases más o menos célebres tenidas por verdades eternas, o sea, principios infalibles de alguna escuela de pensamiento. En efecto, son demasiados los argumentos filosóficos que se ajustan al siguiente molde: "X está equivocado, porque lo que dice contradice lo que escribió el Maestro Y", o bien "El X-ismo es falso, porque sus tesis son incompatibles con las proposiciones fundamentales del Y-ismo". Los dogmáticos —antiguos y modernos, fuera y dentro de la profesión científica, maliciosos o no— obran de esta manera aun cuando no desean convalidar

creencias que simplemente no pueden ser comprobadas, sea empíricamente, sea racionalmente. Porque "dogma" es, por definición, toda opinión no confirmada de la que no se exige verificación porque se la supone verdadera y, más aún, se la supone fuente de verdades ordinarias.

Otro criterio de verdad igualmente difundido ha sido la evidencia. Según esta opinión, verdadero es aquello que parece aceptable a primera vista, sin examen ulterior: aquello, en suma, que se intuye. Así, Aristóteles[3] afirmaba que la intuición "aprehende las premisas primarias" de todo discurso y es por ello "la fuente que origina el conocimiento científico". No sólo Bergson, Husserl y muchos otros intuicionistas e irracionalistas han compartido la opinión de que las esencias pueden captarse sin más: también el racionalismo ingenuo, tal como el que sostenía Descartes, afirma que hay principios evidentes que, lejos de tener que someterse a prueba alguna, son la piedra de toque de toda otra proposición, sea formal o fáctica.

Finalmente, otros han favorecido las "verdades vitales" (o las "mentiras vitales"), esto es, las afirmaciones que se creen o no por conveniencia, independientemente de su fundamento racional y/o empírico. Es el caso de Nietzsche y los pragmatistas posteriores, todos los cuales han exagerado el indudable valor instrumental del conocimiento fáctico, al punto de afirmar que "la posesión de la verdad, lejos de ser [...] un fin en sí, es sólo un medio

[3] Aristóteles, *Analíticos Posteriores*, Libro II, cap. XIX, 110 b.

preliminar para alcanzar otras satisfacciones vitales",[4] de donde "verdadero" es sinónimo de "útil".

Pregúntese a un científico si cree que tiene derecho a suscribir una afirmación en el campo de la ciencia tan sólo porque le guste, o porque la considere un dogma inexpugnable, o porque a él le parezca evidente, o porque la encuentre conveniente. Probablemente conteste más o menos así: ninguno de esos presuntos criterios de verdad garantiza la *objetividad*, y el conocimiento objetivo es la finalidad de la investigación científica. Lo que se acepta sólo por gusto, o por autoridad, o por parecer evidente (habitual), o por conveniencia, no es sino creencia u opinión, pero no es conocimiento científico. El conocimiento científico es a veces desagradable, a menudo contradice a los clásicos (sobre todo si es nuevo), en ocasiones tortura al sentido común y humilla a la intuición: por último, puede ser conveniente para algunos y no para otros. En cambio, aquello que caracteriza al conocimiento científico es su *verificabilidad*: siempre es susceptible de ser verificado (confirmado o disconfirmado).

2. *Veracidad y verificabilidad*

Obsérvese que no pretendemos que el conocimiento científico, por contraste con el ordinario, el

---

[4] W. James, *Pragmatism* (Nueva York, Meridian Books, 1935), p. 134.

tecnológico o el filosófico, sea *verdadero*. Ciertamente lo es con frecuencia, y siempre intenta serlo más y más. Pero la veracidad, que es un objetivo, no caracteriza el conocimiento científico de manera tan inequívoca como el modo, medio o método por el cual la investigación científica plantea problemas y pone a prueba las soluciones propuestas.

En ocasiones, puede alcanzarse una verdad con sólo consultar un texto. Los propios científicos recurren a menudo a un argumento de autoridad atenuada: lo hacen siempre que emplean datos (empíricos o formales) obtenidos por otros investigadores —cosa que no pueden dejar de hacer, pues la ciencia moderna es, cada vez más, una empresa social—. Pero, por grande que sea la autoridad que se atribuye a una fuente, jamás se la considera infalible: si se aceptan sus datos, es sólo provisionalmente y porque se presume que han sido obtenidos con procedimientos que concuerdan con el método científico, de manera que son reproducibles por quienquiera que se disponga a aplicar tales procedimientos. En otras palabras: un dato será considerado verdadero hasta cierto punto, siempre que pueda ser confirmado de manera compatible con los cánones del método científico.

En consecuencia, para que un trozo del saber merezca ser llamado "científico", no basta —ni siquiera es necesario— que sea verdadero. Debemos saber, en cambio, *cómo* hemos llegado a saber, o a presumir, que el enunciado en cuestión es verdadero: debemos ser capaces de enumerar las operaciones (empíricas o racionales) por las cuales es

verificable (confirmable o disconfirmable) de una manera objetiva al menos en principio. Ésta no es sino una cuestión de nombres: quienes no deseen que se exija la verificabilidad del conocimiento deben abstenerse de llamar "científicas" a sus propias creencias, aun cuando lleven bonitos nombres compuestos con raíces griegas. Se les invita cortésmente a bautizarlas con nombres más impresionantes, tales como "reveladas", "evidentes", "absolutas", "vitales", "necesarias para la salud del Estado", "indispensables para la victoria del Partido", etcétera.

Ahora bien, para verificar un enunciado —porque las proposiciones, y no los hechos, son verdaderas o falsas y pueden, por consiguiente, ser verificadas— no basta la contemplación y ni siquiera el análisis. Comprobamos nuestras afirmaciones confrontándolas con otros enunciados. El enunciado confirmatorio (o disconfirmatorio), que puede llamarse el *verificans*, dependerá del conocimiento disponible y de la naturaleza de la proposición dada, la que puede llamarse *verificandum*. Los enunciados confirmatorios serán enunciados referentes a la experiencia si lo que se somete a prueba es una afirmación fáctica, esto es, un enunciado acerca de hechos, sean experimentados o no. Observemos, de pasada, que el científico tiene todo el derecho de especular acerca de hechos inexperienciales, esto es, hechos que en una etapa del desarrollo del conocimiento están más allá del alcance de la experiencia humana; pero entonces está obligado a señalar las experiencias que

permiten inferir tales hechos inobservados o aun inobservables; vale decir, tiene la obligación de anclar sus enunciados fácticos en experiencias conectadas de alguna manera con los hechos transempíricos que supone. Baste recordar la historia de unos pocos inobservables distinguidos: la otra cara de la Luna, las ondas luminosas, los átomos, la conciencia, la lucha de clases y la opinión pública.

En cambio, si lo que se ha verificado no es una proposición referente al mundo exterior, sino un enunciado acerca del comportamiento de signos (tal como, p. ej., '2 + 3 = 5'), entonces los enunciados confirmatorios serán definiciones, axiomas y reglas que se adoptan por una razón cualquiera (p. ej., porque son fecundas en la organización de los conceptos disponibles y en la elaboracion de nuevos conceptos). En efecto, la verificación de afirmaciones pertenecientes al dominio de las formas (lógica y matemática) no requiere otro instrumento material que el cerebro; sólo la verdad fáctica —como en el caso de "la Tierra es redonda"— requiere la observación o el experimento.

Resumiendo: la verificación de enunciados formales sólo incluye operaciones racionales, en tanto que las proposiciones que comunican información acerca de la naturaleza o de la sociedad han de ponerse a prueba por ciertos procedimientos empíricos, tales como el recuento o la medición. Pues, aunque el conocimiento de los hechos no proviene de la experiencia pura —por ser la teoría un componente indispensable de la recolec-

ción de informaciones fácticas— no hay otra manera de verificar nuestras sospechas que recurrir a la experiencia, tanto "pasiva" como activa.

3. Las proposiciones generales verificables: hipótesis científicas

La descripción que antecede satisfará, probablemente, a cualquier científico contemporáneo que reflexione sobre su propia actividad. Pero no resolverá la cuestión para el metacientífico o epistemólogo, para quien los procedimientos, las normas y a veces hasta los resultados de la ciencia son otros tantos problemas. En efecto, el metacientífico no puede dejar de preguntarse cuáles son las afirmaciones verificables, cómo se llega a afirmarlas, cómo se las comprueba y en qué condiciones puede decirse que han sido confirmadas. Tratemos de esbozar una respuesta a estas preguntas.

En primer lugar, si hemos de tratar el problema de la verificación, debemos averiguar *qué* se puede verificar, ya que no toda afirmación —ni siquiera toda afirmación significativa— es verificable. Así, por ejemplo, las definiciones nominales —tales como "América es el continente situado al oeste de Europa"— se aceptan o rechazan sobre la base del gusto, de la conveniencia, etc., pero no pueden verificarse, y ello simplemente porque no son verdaderas ni falsas. Por ejemplo, si convenimos en llamar "norte-sur" a la dirección que normalmente toma la aguja de una brújula, semejan-

te nombre puede gustarnos o no, pero es inverificable: no es sino un nombre, no se funda sobre elemento de prueba alguno, y ninguna operación podría confirmarlo o disconfirmarlo. En cambio, lo que puede confirmarse o disconfirmarse es una afirmación fáctica que contenga a ese término tal como "la 5ª Avenida corre de sur a norte". La verificación de esta afirmación es posible, y puede hacerse con ayuda de una brújula.

No sólo las definiciones nominales, sino también las afirmaciones acerca de fenómenos sobrenaturales son inverificables, puesto que por definición trascienden todo cuanto está a nuestro alcance, y no se las puede poner a prueba con ayuda de la lógica ni de la matemática. Las afirmaciones acerca de la sobrenaturaleza son inverificables no porque no se refieran a hechos —pues a veces pretenden hacerlo—, sino porque no se dispone de *método* alguno mediante el cual se pueda decidir cuál es su valor de verdad. En cambio, muchas de ellas son perfectamente significativas para quien se tome el trabajo de ubicarlas en su contexto sin pretender reducirlas, por ejemplo, a conceptos científicos. La verificación torna más exacto el significado, pero no produce significado alguno. Más bien al contrario, la posesión de un significado determinado es una condición necesaria para que una proposición sea verificable. Pues ¿cómo habríamos de disponernos a comprobar lo que no entendemos?

Ahora bien, los enunciados verificables son de muchas clases. Hay proposiciones singulares, ta-

les como "Este trozo de hierro está caliente"; particulares o existenciales, tales como "Algunos trozos de hierro están calientes"; proposiciones universales, como "Todos los trozos de hierro están calientes" (que es verificablemente falsa). Hay, además, enunciados de leyes, tales como "Todos los metales se dilatan con el calor" (o mejor, "Para todo x, si x es un trozo de metal que se calienta, entonces x se dilata"). Las proposiciones singulares y particulares pueden verificarse a menudo de manera inmediata, con la sola ayuda de los sentidos o, eventualmente, con el auxilio de instrumentos que amplíen su alcance; pero otras veces exigen operaciones complejas que implican enunciados de leyes y cálculos matemáticos, como en el caso de "La distancia media entre la Tierra y el Sol es de unos 1.500 millones de kilómetros".

Cuando un enunciado verificable posee un grado de generalidad suficiente, habitualmente se lo llama *hipótesis* científica. O, lo que es equivalente, cuando una proposición general (particular o universal) puede verificarse sólo de manera *indirecta* —esto es, por el examen de algunas de sus consecuencias— es conveniente llamarla "hipótesis científica". Por ejemplo, "Todos los trozos de hierro se dilatan con el calor", y, *a fortiori*, "Todos los metales se dilatan con el calor", son hipótesis científicas: son puntos de partida de raciocinios y, por ser generales, sólo pueden ser confirmados poniendo a prueba sus consecuencias particulares, esto es, probando enunciados referentes a muestras específicas de metal.

Solía creerse que el discurso científico no incluía elementos hipotéticos sino tan sólo hechos y, sobre todo, lo que en inglés se denomina *hard facts*. Ahora se comprende que el núcleo de toda teoría científica es un conjunto de hipótesis verificables. Las hipótesis científicas son, por una parte, remates de cadenas inferenciales no demostrativas (analógicas o inductivas) más o menos oscuras; por otra parte, son puntos de partida de cadenas deductivas cuyos últimos eslabones —los más próximos a los sentidos, en el caso de la ciencia fáctica— deben pasar la prueba de la experiencia.

Más aún: habitualmente se concuerda en que debiera llamarse "hipótesis" no sólo a las conjeturas de ensayo, sino también a las suposiciones razonablemente confirmadas o establecidas, pues probablemente no hay enunciados fácticos generales perfectos. La experiencia ha sugerido adoptar este sentido de la palabra "hipótesis". Considérese, por ejemplo, la ley de Newton de la gravedad, que ha sido confirmada en *casi* todos los casos con una precisión asombrosa. Tenemos dos razones para llamarla "hipótesis": la primera es que ha pasado la prueba sólo un número finito de veces; la segunda, es que hemos terminado por aprender que incluso ese célebre enunciado de ley es tan sólo una primera aproximación de un enunciado más exacto incluido en la teoría general de la relatividad, que tampoco es probable que sea definitiva.

4. *El método científico, ¿*ars inveniendi*?*

Hemos convenido en que un enunciado fáctico general susceptible de ser verificado puede llamarse "hipótesis", lo que suena más respetable que corazonada, sospecha, conjetura, suposición o presunción, y es también más adecuado que estos términos, ya que según su etimología "hipótesis" es punto de partida, que ciertamente lo es una vez que se ha dado con ella. Abordemos ahora el segundo problema que nos propusimos, a saber: ¿existe una técnica infalible para inventar hipótesis científicas que sean probablemente verdaderas? En otras palabras: ¿existe un método, en el sentido cartesiano de conjunto de "reglas ciertas y fáciles" que nos conduzca a enunciar verdades fácticas de gran extensión?

Muchos hombres, en el curso de muchos siglos, han creído en la posibilidad de descubrir la técnica del descubrimiento, y de inventar la técnica de la invención. Fue fácil bautizar al niño no nacido, y se lo hizo con el nombre de *ars inveniendi*. Pero semejante arte jamás fue inventado. Lo que es más, podría argüirse que jamás se lo inventará, a menos que se modifique radicalmente la definición de "ciencia"; en efecto, el conocimiento científico, por oposición a la sabiduría revelada, es esencialmente *falible*, esto es, susceptible de ser parcial o aun totalmente refutado. La falibilidad del conocimiento científico, y por consiguiente la imposibilidad de establecer reglas de oro que nos

conduzcan derechamente a verdades finales, no es sino el complemento de aquella verificabilidad que habíamos encontrado en el núcleo de la ciencia.

Vale decir, no hay reglas infalibles que *garanticen* por anticipado el descubrimiento de nuevos hechos y la invención de nuevas teorías, asegurando así la fecundidad de la investigación científica: la certidumbre debe buscarse tan sólo en las ciencias formales. ¿Significa esto que la investigación científica es errática e ilegal, y por consiguiente que los científicos lo esperan todo de la intuición o de la iluminación? Tal es la moraleja que algunos científicos y filósofos eminentes han extraído de la inexistencia de leyes que nos *aseguren* contra la infertilidad y el error. Por ejemplo, Bridgman —el expositor del operacionismo— ha negado la existencia del método científico, sosteniendo que "la ciencia es lo que hacen los científicos, y hay tantos métodos científicos como hombres de ciencia".[5]

Es verdad que en ciencia no hay caminos reales, que la investigación se abre camino en la selva de los hechos, y que los científicos sobresalientes elaboran su propio estilo de pesquisa. Sin embargo, esto no debe hacernos desesperar de la posibilidad de descubrir pautas, normalmente satisfactorias, de *plantear problemas y poner a prueba hipótesis*. Los científicos que van en pos de la verdad no se comportan ni como soldados que

---

[5] P. W, Bridgman, *Reflections of a Physicist* (Nueva York, Philosophical Library, 1955), p. 83.

cumplen obedientemente las reglas de la ordenanza (opiniones de Bacon y Descartes), ni como los caballeros de Mark Twain, que cabalgaban en cualquier dirección para llegar a Tierra Santa (opinión de Bridgman). No hay avenidas hechas en ciencia, pero hay, en cambio, una brújula mediante la cual a menudo es posible estimar si se está sobre una huella promisoria. Esta brújula es el método científico, que no produce automáticamente el saber, pero que nos evita perdernos en el caos aparente de los fenómenos, aunque sólo sea porque nos indica cómo *no* plantear los problemas y cómo *no* sucumbir al embrujo de nuestros prejuicios predilectos.

La investigación no es errática sino metódica; sólo que no hay una sola manera de sugerir hipótesis, sino muchas maneras: las hipótesis no se nos imponen por la fuerza de los hechos, sino que son *inventadas* para dar cuenta de los hechos. Es verdad que la invención no es ilegal, sino que sigue ciertas pautas; pero éstas son psicológicas antes que lógicas, son peculiares de los diversos tipos intelectuales, y por añadidura las conocemos poco porque apenas se las investiga. Hay, ciertamente, reglas que *facilitan* la invención científica, y en especial la formulación de hipótesis; entre ellas figuran las siguientes: el sistemático reordenamiento de los datos, la supresión imaginaria de factores con el fin de descubrir las variables pertinentes, el obstinado cambio de representación en busca de analogías fructíferas. Sin embargo, las reglas que favorecen o entorpecen el trabajo cien-

tífico no son de oro sino plásticas; más aún, el investigador rara vez tiene conciencia del camino que ha tomado para formular sus hipótesis. Por esto la investigación científica puede planearse a grandes líneas y no en detalle, y aun menos puede ser regimentada.

Algunas hipótesis se formulan por vía inductiva, esto es, como generalizaciones sobre la base de la observación de un puñado de casos particulares. Pero la inducción dista de ser la única o siquiera la principal de las vías que conducen a formular enunciados generales verificables. Otras veces, el científico opera por analogía; por ejemplo, la teoría ondulatoria de la luz le fue sugerida a Huyghens (1690) por una comparación con las olas.[6] En algunos casos el principio heurístico es una analogía matemática; así, por ejemplo, Maxwell (1873) predijo la existencia de ondas electromagnéticas sobre la base de una analogía formal entre sus ecuaciones del campo y la conocida ecuación de las ondas elásticas.[7] Ocasionalmente, el investigador es guiado por consideraciones filosóficas; así fue como procedió Oersted (1820); buscó deliberadamente una conexión entre la electricidad y el magnetismo, obrando sobre la base de la convicción *a priori* de que la estructura de todo cuanto existe es polar, y que todas las

---

[6] C. Huyghens, *Traité de la lumière* (París, Gauthier-Villers, 1920), p. 5.

[7] J. C. Maxwell, *A Treatise of Electricity and Magnetism*, 3ª ed. (Oxford, University Press, 1937), II, pp. 434 y ss.

"fuerzas" de la naturaleza están conectadas orgánicamente entre sí.[8] La convicción filosófica de que la complejidad de la naturaleza es ilimitada llevó a Bohm a especular sobre un nivel subcuántico, fundándose en una analogía con el movimiento browniano clásico.[9] Ni siquiera la fantasía teológica ha dejado de contribuir, aunque por cierto en mínima medida; recuérdese el principio de la mínima acción, de Maupertuis (1747), formulado en la creencia de que el Creador lo había dispuesto todo de la manera más económica posible.

A las hipótesis científicas se llega, en suma, de muchas maneras: hay muchos principios heurísticos, y el único invariante es el requisito de verificabilidad. La inducción, la analogía y la deducción de suposiciones extracientíficas (p. ej., filosóficas) proveen puntos de partida que deben ser elaborados y probados.

5. El método científico, técnica de planteo y comprobación

Los especialistas científicos habitualmente no se interesan por el problema de la génesis de las

---

[8] Véase, p ej., S. F. Mason, *A History of the Sciences* (Londres, Routledge & Kegan Paul, 1953), p. 386.

[9] D. Bohm, "A Proposed Explanation of Quantum Theory in Terms of Hidden Variables at a Sub Quantum Mechanical Level", en Colston Papers (Londres, Butterworths Scientific Publications, 1957), IX, p. 33.

hipótesis científicas; esta cuestión es de competencia de las diversas ciencias de la ciencia. El proceso que conduce a la enunciación de una hipótesis científica puede estudiarse en diversos niveles: el lógico, el psicológico y el sociológico. El lógico se interesará por la inferencia plausible como conexión inversa (no deductiva) entre proposiciones singulares y generales. El psicólogo investigará la etapa de la "iluminación" o relámpago en el proceso de la resolución de los problemas, etapa en que se produce la síntesis de elementos anteriormente inconexos; también se propondrá estudiar fenómenos tales como los estímulos e inhibiciones que caracterizan el trabajo en equipo. El sociólogo inquirirá por qué determinada estructura social favorece ciertas clases de hipótesis mientras desalienta otras.

El metodólogo, en cambio, no se ocupará de la génesis de las hipótesis, sino del *planteo* de los problemas que las hipótesis intentan resolver, y de su *comprobación*. El origen del nexo entre el planteo y la comprobación —esto es, el surgimiento de la hipótesis— lo deja a otros especialistas. El motivo es, nuevamente, una cuestión de nombres: lo que hoy se llama "método científico" no es ya una lista de recetas para dar con las respuestas correctas a las preguntas científicas, sino el conjunto de procedimientos por los cuales a) se plantean los problemas científicos, y b) se ponen a prueba las hipótesis científicas.

El estudio del método científico es, en una palabra, la teoría de la investigación. Esta teoría

es descriptiva en la medida en que descubre pautas en la investigación científica (y aquí interviene la historia de la ciencia, como proveedora de ejemplos). La metodología es normativa en la medida en que muestra cuáles son las reglas de procedimiento que pueden aumentar la probabilidad de que el trabajo sea fecundo. Pero las reglas discernibles en la práctica científica exitosa son perfectibles: no son cánones intocables porque no garantizan la obtención de la verdad; pero, en cambio, facilitan la detección de errores.

Si la hipótesis que ha de ser puesta a prueba se refiere a objetos ideales (números, funciones, figuras, fórmulas lógicas, suposiciones filosóficas, etc.), su verificación consistirá en la prueba de su coherencia —o incoherencia— con enunciados (postulados, definiciones, etc.) previamente aceptados. En este caso, la confirmación puede ser una demostración definitiva. En cambio, si el enunciado en cuestión se refiere (de manera significativa) a la naturaleza o a la sociedad, puede ocurrir, o bien que podamos averiguar su valor de verdad con la sola ayuda de la razón, o que debamos recurrir, además, a la experiencia.

El análisis lógico basta cuando el enunciado que se pone a prueba es de alguno de los siguientes tipos: a) una simple tautología, o sea, un enunciado verdadero en virtud de su sola forma, independientemente de su contenido (como en el caso de "El agua moja o no moja"); b) una definición, o equivalencia entre dos grupos de términos (como en el caso de "Los seres vivos se alimentan, crecen

y se reproducen"); c) una consecuencia de enunciados fácticos que poseen una extensión o alcance mayor (como ocurre cuando se deduce el "principio" de la palanca, de la ley de conservación de la energía). Vale decir, el análisis lógico y matemático comprobará la validez de los enunciados (hipótesis) que son analíticos en determinado contexto. Muchos enunciados no son intrínsecamente analíticos: su analiticidad es relativa o contextual, como lo demuestra el hecho de que esta propiedad puede perderse si se estrecha o amplía el contexto, o si se reagrupan los enunciados de la teoría correspondiente, de manera tal que los antiguos teoremas se conviertan en postulados y viceversa.

Vale decir, la mera *referencia* a los hechos no basta para decidir qué herramienta, si el análisis o la experiencia, ha de emplearse para convalidar una proposición: hay que empezar por determinar su *status* y estructura lógica. En consecuencia, *el análisis lógico (tanto sintáctico como semántico) es la primera operación que debiera emprenderse al comprobar las hipótesis científicas*, sean fácticas o no. Esta norma debiera considerarse como una regla del método científico.

Los enunciados fácticos no analíticos —esto es, las proposiciones referentes a hechos pero indecidibles con la sola ayuda de la lógica— tendrán que concordar con los datos empíricos o adaptarse a ellos. Esta norma, que distaba de ser obvia antes del siglo XVII, y que contradice tanto el apriorismo escolástico como el racionalismo cartesiano, es la segunda regla del método científico. Podemos

enunciarla de la siguiente manera: *El método científico, aplicado a la comprobación de afirmaciones informativas, se reduce al método experimental.*

6. El método experimental

La experimentación involucra la modificación deliberada de algunos factores, es decir, la sujeción del objeto de experimentación a estímulos controlados. Pero lo que habitualmente se llama "método experimental" no involucra necesariamente experimentos en el sentido estricto del término, y puede aplicarse fuera del laboratorio. Así, por ejemplo, la astronomía no experimenta con cuerpos celestes (por el momento) pero es una ciencia empírica porque se funda sobre la física, que aplica el método experimental. En lugar de elaborar una definición de este término, veamos cómo funcionó en un caso famoso tan conocido que casi siempre se entiende mal.

Adams y Le Verrier descubrieron el planeta Neptuno procediendo de una manera que es típica de la ciencia moderna. Sin embargo, no ejecutaron un solo experimento; ni siquiera partieron de "hechos sólidos". En efecto, el problema que se plantearon fue el de explicar ciertas irregularidades halladas en el movimiento de los planetas exteriores (a la Tierra); pero estas irregularidades no eran fenómenos observables: consistían en discrepancias entre las órbitas *observadas* y las *calcula-*

das. El hecho que debían explicar no era un conjunto de datos de los sentidos, sino un conflicto entre datos empíricos y consecuencias deducidas de los principios de la mecánica celeste.

La hipótesis que propusieron para explicar la discrepancia fue que un planeta transuraniano inobservado perturbaba el movimiento de los planetas exteriores entonces conocidos. También podrían haber imaginado que la ley de Newton de la gravitación falla a grandes distancias, pero esto era apenas concebible en una época en que la *Weltanschauung* prevaleciente entre los científicos incluía una fe dogmática en la física newtoniana. De esta hipótesis, unida a los principios aceptados de la mecánica celeste y ciertas suposiciones específicas (referentes, entre otras, al plano de la órbita), Adams y Le Verrier dedujeron consecuencias observables con la sola ayuda de la lógica y de la matemática: predijeron el lugar en que se encontraría el "nuevo" planeta en tal y cual noche. La observación del cielo y el descubrimiento del planeta en el lugar y el momento predichos no fueron sino el último eslabón de un largo proceso por el cual se probaron conjuntamente varias hipótesis.

No es fácil decidir si una hipótesis concuerda con los hechos. En primer lugar, la verificación empírica rara vez puede determinar cuál de los componentes de una teoría dada ha sido confirmado o disconfirmado; habitualmente se prueban sistemas de proposiciones antes que enunciados aislados. Pero la principal dificultad proviene de la generalidad de las hipótesis científicas. La hipóte-

sis de Adams y Le Verrier era general, aun cuando ello no es aparente a primera vista: tácitamente habían supuesto que el planeta existía *en todo momento* dentro de un largo lapso; y comprobaron la hipótesis tan sólo para unos pocos breves intervalos de tiempo. En cambio, las proposiciones fácticas singulares no son tan difíciles de probar. Así, p. ej., no es difícil comprobar si "El Sr. Pérez, que es obeso, es cardíaco": bastan una balanza y un estetoscopio. Lo difícil de comprobar son las proposiciones fácticas generales, esto es, los enunciados referentes a *clases* de hechos y no a hechos singulares. La razón es sencilla; no hay hechos generales, sino tan sólo hechos singulares; por consiguiente, la frase "adecuación de las ideas a los hechos" está fuera de la cuestión en lo que respecta a las hipótesis científicas.

Supongamos que se sugiere la hipótesis "Los obesos son cardíacos", sea por la observación de cierto número de correlaciones entre la obesidad y las enfermedades del corazón (esto es, por inducción estadística), sea sobre la base del estudio de la función del corazón en la circulación (esto es, por deducción). El enunciado general "Los obesos son cardíacos" no se refiere solamente a nuestros conocidos, sino a todos los gordos del mundo; por consiguiente, no podemos esperar verificarlo *directamente* (esto es, por el examen de un inexistente "gordo general") ni *exhaustivamente* (auscultando a todos los seres humanos presentes, pasados y futuros). La metodología nos dice cómo debemos proceder; en este caso, examinaremos su-

cesivamente los miembros de una muestra suficientemente numerosa de personas obesas. Vale decir, probamos una consecuencia particular de nuestra suposición general. Ésta es una tercera máxima del método científico: *Obsérvense singulares en busca de elementos de prueba de universales.*

Hasta aquí todo parece sencillo; pero los problemas relacionados con la prueba real distan de ser triviales, y algunos de ellos no han sido resueltos satisfactoriamente. Debemos recurrir a las técnicas del planteo de problemas de este tipo, es decir, a las técnicas de diseño de los procedimientos empíricos adecuados. Esta técnica nos aconseja comenzar por decidir lo que hemos de entender por "obeso" y por "cardíaco", lo que no es en modo alguno tarea sencilla, ya que el umbral de obesidad es en gran medida convencional. O sea, debemos empezar por determinar el exacto sentido de nuestra pregunta. Y ésta es una cuarta regla del método científico, a saber: *Formúlense preguntas precisas.*

Luego procederemos a elegir la técnica experimental (clase de balanza, tipo de examen de corazón, etc.) y la manera de registrar datos y de ordenarlos. Además, debemos decidir el tamaño de la muestra que habremos de observar y la técnica para escoger sus miembros, con el fin de asegurar que será una fiel representante de la población total. Sólo una vez realizadas estas operaciones preliminares podremos visitar al señor Pérez y a los demás miembros de la muestra, con el fin de reu-

nir datos. Y aquí se nos muestra una quinta regla del método científico: *La recolección y el análisis de datos deben hacerse conforme a las reglas de la estadística.*

Después que los datos han sido reunidos, clasificados y analizados, el equipo que tiene a su cargo la investigación podrá realizar una inferencia estadística, concluyendo que "El N % de los obesos son cardíacos". Más aún, habrá que estimar el error probable de esta afirmación.

Obsérvese que la hipótesis que había motivado nuestra investigación era un enunciado universal de la forma "Para todo x, si x es F, entonces x es G". Por otro lado, el resultado de la investigación es un enunciado estadístico, a saber, "De la clase de las personas obesas, una subclase que comprende el X % de la población examinada está compuesta por cardíacos". Esto es, nuestra hipótesis de trabajo ha sido corregida. ¿Debemos contentarnos con esta respuesta? Nos gustaría formular otras preguntas: deseamos entender la ley que hemos hallado, nos gustaría deducirla de las leyes de la fisiología humana. Y aquí se aplica una sexta regla del método científico, a saber: *No existen respuestas definitivas, y ello simplemente porque no existen preguntas finales.*

7. *Métodos teóricos*

Toda ciencia fáctica especial elabora sus propias técnicas de verificación; entre ellas, las técni-

cas de medición son típicas de la ciencia moderna. Pero en todos los casos estas técnicas, por diferentes que sean, no constituyen fines en sí mismas: todas ellas sirven para contrastar ciertas ideas con ciertos hechos por la vía de la experiencia. O, si se prefiere, el objetivo de las técnicas de verificación es probar enunciados referentes a hechos por vía del examen de proposiciones referentes a la experiencia (y, en particular, al experimento). Éste es el motivo por el cual los experimentadores no tienen por qué construir cada uno de sus aparatos e instrumentos, pero deben en cambio diseñarlos y/o usarlos a fin de poner a prueba ciertas afirmaciones. Las técnicas especiales, por importantes que sean, no son sino etapas de la aplicación del método experimental, que no es otra cosa que el método científico en relación con la ciencia fáctica; y la ciencia, por fáctica que sea, no es un montón de hechos sino un sistema de ideas.

En el párrafo anterior ejemplificamos el método experimental analizando el proceso de verificación que requeriría el enunciado "Los obesos son cardíacos"; encontramos que esta hipótesis requería una precisión cuantitativa y, después de una investigación imaginaria, adoptamos, en su lugar, cierta generalización empírica del tipo de los enunciados estadísticos. Ahora bien: las generalizaciones empíricas, tan caras a Aristóteles y a Bacon, aun cuando se las formule en términos estadísticos, no son distintivas de la ciencia moderna. El tipo de hipótesis característico de la ciencia moderna no es el de los enunciados descriptivos

aislados, cuya función principal es resumir experiencias. Lo peculiar de la ciencia moderna es que consiste en su mayor parte en *teorías explicativas*, es decir, en sistemas de proposiciones que pueden clasificarse en: principios, leyes, definiciones, etcétera, y que están vinculadas entre sí mediante conectivas lógicas (tales como "y", "o", "si... entonces", etcétera).

Las teorías dan cuenta de los hechos no sólo describiéndolos de manera más o menos exacta, sino también proveyendo modelos conceptuales de los hechos, en cuyos términos puede explicarse y predecirse, al menos en principio, cada uno de los hechos de una clase. Las posibilidades de una hipótesis científica no se advierten por entero antes de incorporarlas en una teoría; y es sólo entonces cuando puede encontrársele varios soportes. Al sumergirse en una teoría, el enunciado dado es apoyado —o aplastado— por toda la masa del saber disponible; permaneciendo aislado es difícil de confirmar y de refutar y, sobre todo, sigue sin ser entendido.

La conversión de las generalizaciones empíricas en leyes teóricas implica trascender la esfera de los fenómenos y el lenguaje observacional: ya no se trata de hacer afirmaciones acerca de hechos observables, sino de adivinar su "mecanismo" interno (el que, desde luego, no tiene por qué ser mecánico). Supóngase que un psicólogo desea estudiar las correlaciones entre cierto estímulo observable S y cierta conducta observable R, que —a modo de ensayo— considera como la respuesta al

estímulo dado. Si, después de una sucesión de experimentos, llegara a confirmar su hipótesis de trabajo y deseara trascender las fronteras de la psicología fenomenista, intentaría elaborar, digamos, un modelo neurofisiológico que explicara el nexo *S-R* en términos fisiológicos. No es tarea fácil: el psicólogo tiene que inventar diversas hipótesis acerca de otros tantos canales nerviosos posibles que conecten los hechos observables extremos, *S* y *R*. Análogamente, los físicos atómicos imaginan diversos mecanismos ocultos que conectan los fenómenos macroscópicos con su soporte microscópico.

Pero nuestro psicólogo no andará del todo a tientas: podrá probar si su conexión concuerda con algunos de los esquemas conocidos, por ejemplo la ley de Hebb de los refuerzos de las conexiones interneuronales. Cada una de sus hipótesis —sea que consistan en suponer que interviene un reflejo innato o condicionado— tendrá que especificar el aparato receptor, el nervio aferente, la estación central, el nervio eferente, el órgano receptor, etc. Más aún, sus varias hipótesis de trabajo tendrán que ser compatibles con el saber más firmemente establecido (aunque no inamovible), y tendrán que ser puestas a prueba mediante técnicas especiales (excitación o destrucción de nervios, registro de impulsos nerviosos, medición del consumo de glucosa, etc.). Vale la pena emprender esta difícil tarea: la eventual confirmación de una de las hipótesis puestas a prueba no sólo explicará el nexo *S-R* dado, sino que también lo ubicará en su con-

texto; además, apoyará la hipótesis misma de que tal nexo no es accidental. Pues, aunque suene a paradoja, un enunciado fáctico es tanto más fidedigno cuanto mejor está apoyado por consideraciones teóricas.

Es importante advertir, en efecto, que la experiencia dista de ser el único juez de las teorías fácticas, o siquiera el último. Las teorías se contrastan con los hechos y con *otras teorías*. Por ejemplo, una de las pruebas de la generalización de una teoría dada es averiguar si la nueva teoría se reduce a la vieja dentro de un cierto dominio, de modo tal que cubra por lo menos el mismo grupo de hechos. Más aún, el *grado de sustentación* o apoyo de las teorías no es idéntico a su *grado de confirmación*. Las teorías no se construyen *ex nihilo*, sino sobre ciertas bases: éstas las sostienen antes y después de la prueba; la prueba misma, si tiene éxito, provee los apoyos restantes de la teoría y fija su grado de confirmación. Aun así el grado de confirmación de una teoría no basta para determinar la verosimilitud de la misma.

8. En qué se apoya una hipótesis científica

Una hipótesis de contenido fáctico no sólo es sostenida por la confirmación empírica de cierto número de sus consecuencias particulares (p. ej., predicciones). Las hipótesis científicas están incorporadas en teorías o tienden a incorporarse en ellas; y las teorías están relacionadas entre sí,

constituyendo la totalidad de ellas la cultura intelectual. Por esto, no debiera sorprender que las hipótesis científicas tengan soportes no sólo científicos sino también extracientíficos; los primeros son empíricos y racionales, los últimos son psicológicos y culturales. Expliquémonos.

Cuanto más numerosos sean los hechos que confirman una hipótesis, cuanto mayor sea la precisión con que ella reconstruye los hechos, y cuanto más vastos sean los nuevos territorios que ayuda a explorar, tanto más firme será nuestra creencia en ella, esto es, tanto mayor será la plausibilidad o verosimilitud que le asignemos. Esto es, esquemáticamente dicho, lo que se entiende por el *soporte empírico* de las hipótesis fácticas. Pero la experiencia disponible no puede ser considerada como inapelable: en primer lugar, porque nuevas experiencias pueden mostrar la necesidad de un remiendo; en segundo término, porque la experiencia científica no es pura, sino interpretada, y toda interpretación se hace en términos de teorías, motivo por el cual la primera reacción de los científicos experimentados ante informaciones sobre hechos que parecerían trastornar teorías establecidas, es de escepticismo.

Cuanto más estrecho sea el acuerdo de la hipótesis en cuestión con el conocimiento disponible del mismo orden, tanto más firme es nuestra creencia en ella; semejante concordancia es particularmente valiosa cuando consiste en una compatibilidad con enunciados de leyes. Esto es lo que hemos designado con el nombre de *soporte* racio-

nal de las hipótesis fácticas. Éste es, dicho sea de paso, el motivo por el cual la mayoría de los científicos desconfían de los informes acerca de la llamada percepción extrasensorial, porque los llamados fenómenos *psi* contradicen el cuerpo de hipótesis psicológicas y fisiológicas bien establecidas. En resumen, las teorías científicas deben adecuarse, sin duda, a los hechos, pero ningún hecho aislado es aceptado en la comunidad de los hechos controlados científicamente a menos que tenga cabida en alguna parte del edificio teórico establecido. Desde luego, el soporte racional no es garantía de verdad; si lo fuera, las teorías fácticas serían invulnerables a la experiencia. Los soportes empíricos y racionales de las hipótesis fácticas son interdependientes.

En cuanto a los soportes extracientíficos de las hipótesis científicas, uno de ellos es de carácter psicológico: influye sobre nuestra elección de las suposiciones y sobre el valor que le asignamos a su concordancia con los hechos. Por ejemplo, los sentimientos estéticos que provocan la simplicidad y la unidad lógica estimulan unas veces y otras obstaculizan la investigación sobre la validez de las teorías. Esto es lo que hemos denominado el *soporte psicológico* de las hipótesis fácticas; a menudo es oscuro, y no sólo está vinculado a características personales, sino también sociales.

Lo que hemos llamado *soporte cultural* de las hipótesis fácticas consiste en su compatibilidad con alguna concepción del mundo y, en particular, con la *Zeitgeist* prevaleciente. Es obvio que tende-

mos a asignar mayor peso a aquellas hipótesis que congenian con nuestro fondo cultural y, en particular con nuestra visión del mundo, que a aquellas hipótesis que lo contradicen. Por ejemplo, la hipótesis de la evolución molecular fue bien recibida en la década de 1970 (y luego confirmada) porque la concepción moderna del mundo es evolucionista.

La función dual del soporte cultural de las conjeturas científicas se advierte con facilidad: por una parte, nos impulsa a poner atención en ciertas clases de hipótesis y hasta interviene en la sugerencia de las mismas; por otra parte, puede impedirnos apreciar otras posibilidades, por lo cual puede constituir un factor de obstinación dogmática. La única manera de minimizar este peligro es cobrar conciencia del hecho de que las hipótesis científicas no crecen en un vacío cultural.

Los soportes empíricos y racionales son objetivos, en el sentido de que en principio son susceptibles de ser sopesados y controlados conforme a patrones precisos y formulables. En cambio, los soportes extracientíficos son, en gran medida, materia de preferencia individual, de grupo o de época; por consiguiente, no debieran ser decisivos en la etapa de la comprobación, por prominentes que sean en la etapa heurística. Es importante que los científicos sean personas cultas, aunque sólo sea para que adviertan la fuerte presión que ejercen los factores psicológicos y culturales sobre la formulación, elección, investigación y credibilidad de las hipótesis fácticas. La presión, para bien o para

mal, es real y nos obliga a tomar partido por una u otra concepción del mundo; es mejor hacerlo conscientemente que inadvertidamente.

La enumeración anterior de los tipos de soporte de las hipótesis científicas no tenía otro propósito que mostrar que el método experimental no agota el proceso que conduce a la aceptación de una suposición fáctica. Este hecho podría invocarse en favor de la tesis de que la investigación científica es un arte.

9. La ciencia: técnica y arte

La investigación científica es legal, pero sus leyes —las reglas del método científico— no son pocas, ni simples, ni infalibles, ni bien conocidas; son, por el contrario, numerosas, complejas, más o menos eficaces, y en parte desconocidas. El arte de formular preguntas y de probar respuestas —esto es, el método científico— es cualquier cosa menos un conjunto de recetas; y menos técnica todavía es la teoría del método científico. La moraleja es inmediata: desconfíese de toda descripción de la vía de la ciencia —y en primer lugar de la presente—, pero no se descuide ninguna. La investigación es una empresa multilateral que requiere el más intenso ejercicio de cada una de las facultades psíquicas, y que exige un concurso de circunstancias sociales favorables; por este motivo, todo testimonio personal, perteneciente a cualquier período, y por parcial que sea, puede echar

alguna luz sobre algún aspecto de la investigación.

A menudo se sostiene que la medicina y otras ciencias aplicadas son artes antes que ciencias, en el sentido de que no pueden ser reducidas a la simple aplicación de un conjunto de reglas que pueden formularse todas explícitamente y que pueden elegirse sin que medie el juicio personal. Sin embargo, en este sentido la física y la matemática también son artes: ¿quién conoce recetas hechas y seguras para encontrar leyes de la naturaleza o para adivinar teoremas? Si "arte" significa una feliz conjunción de experiencia, destreza, imaginación, visión y habilidad para realizar inferencias de tipo no analítico, entonces no sólo son artes la medicina, la pesquisa criminal, la estrategia militar, la política y la publicidad, sino también toda otra disciplina. Por consiguiente, no se trata de si un campo dado de la actividad humana es un arte, sino de si, además, es científico.

La ciencia es ciertamente comunicable; si un cuerpo de conocimiento no es comunicable, entonces por definición no es científico. Pero esto se refiere a los *resultados* de la investigación antes que a las maneras en que éstos se obtienen; la comunicabilidad no implica que el método científico y las técnicas de las diversas ciencias especiales puedan aprenderse en los libros: los procedimientos de la investigación se dominan investigando, y los metacientíficos debieran por ello practicarlos antes de emprender su análisis. No se sabe de obra maestra alguna de la ciencia que haya sido

engendrada por la aplicación consciente y escrupulosa de las reglas conocidas del método científico; la investigación científica es practicada en gran parte como un arte no tanto porque carezca de reglas cuanto porque algunas de ellas se dan por sabidas, y no tanto porque requiera una intuición innata cuanto porque exige una gran variedad de disposiciones intelectuales. Como toda otra experiencia, la investigación puede ser comprendida por otros pero no es íntegramente transferible; hay que pagar por ella el precio de un gran número de errores, y por cierto que al contado. Por consiguiente, los escritos sobre el método científico pueden iluminar el camino de la ciencia, pero no pueden exhibir toda su riqueza y, sobre todo, no son un sustituto de la investigación misma, del mismo modo que ninguna biblioteca sobre botánica puede reemplazar la contemplación de la naturaleza, aunque hace posible que la contemplación sea más provechosa.

10. La pauta de la investigación científica

La variedad de habilidades y de información que exige el tratamiento científico de los problemas ayuda a explicar la extremada división del trabajo prevaleciente en la ciencia contemporánea, en la que encuentra lugar toda capacidad natural y toda habilidad adquirida. Es posible apreciar esta variedad exponiendo la pauta general de la investigación científica. Creo que esta pauta

—o sea, el método científico— es, a grandes líneas, la siguiente:

1. PLANTEO DEL PROBLEMA
   1.1. *Reconocimiento de los hechos:* examen del grupo de hechos, clasificación preliminar y selección de los que probablemente sean relevantes en algún respecto.
   1.2. *Descubrimiento del problema:* hallazgo de la laguna o de la incoherencia en el cuerpo del saber.
   1.3. *Formulación del problema:* planteo de una pregunta que tiene visos de ser la correcta; esto es, reducción del problema a su núcleo significativo, probablemente soluble y probablemente fructífero, con ayuda del conocimiento disponible.

2. CONSTRUCCIÓN DE UN MODELO TEÓRICO
   2.1. *Selección de los factores pertinentes:* invención de suposiciones plausibles relativas a las variables que probablemente son pertinentes.
   2.2. *Invención de las hipótesis centrales y de las suposiciones auxiliares:* propuesta de un conjunto de suposiciones concernientes a los nexos entre las variables pertinentes; p. ej., formulación de enunciados de ley que se espera puedan amoldarse a los hechos observados.
   2.3. *Traducción matemática:* cuando sea posible, traducción de las hipótesis, o de parte

de ellas, a alguno de los lenguajes matemáticos.

3. DEDUCCIÓN DE CONSECUENCIAS PARTICULARES
   3.1. *Búsqueda de soportes racionales:* deducción de consecuencias particulares que pueden haber sido verificadas en el mismo campo o en campos contiguos.
   3.2. *Búsqueda de soportes empíricos:* elaboración de predicciones (o retrodicciones) sobre la base del modelo teórico y de datos empíricos, teniendo en vista técnicas de verificación disponibles o concebibles.

4. PRUEBA DE LAS HIPÓTESIS
   4.1. *Diseño de la prueba:* planeamiento de los medios para poner a prueba las predicciones; diseño de observaciones, mediciones, experimentos y demás operaciones instrumentales.
   4.2. *Ejecución de la prueba:* realización de las operaciones y recolección de datos.
   4.3. *Elaboración de los datos:* clasificación, análisis, evaluación, reducción, etcétera, de los datos empíricos.
   4.4. *Inferencia de la conclusión:* interpretación de los datos elaborados a la luz del modelo teórico.

5. INTRODUCCIÓN DE LAS CONCLUSIONES EN LA TEORÍA
   5.1. *Comparación de las conclusiones con las predicciones:* contraste de los resultados

de la prueba con las consecuencias del modelo teórico, precisando en qué medida éste puede considerarse confirmado o disconfirmado (inferencia probable).

5.2. *Reajuste del modelo:* eventual corrección o aun reemplazo del modelo.

5.3. *Sugerencias acerca del trabajo ulterior:* búsqueda de lagunas o errores en la teoría y/o los procedimientos empíricos, si el modelo ha sido disconfirmado; si ha sido confirmado, examen de posibles extensiones y de posibles consecuencias en otros departamentos del saber.

11. Extensibilidad del método científico

Para elaborar conocimiento fáctico no se conoce mejor camino que el de la ciencia. El método de la ciencia no es, por cierto, seguro; pero es intrínsecamente progresivo, porque es autocorrectivo: exige la continua comprobación de los puntos de partida, y requiere que todo resultado sea considerado como fuente de nuevas preguntas. Llamemos *filosofía científica* a la clase de concepciones filosóficas que aceptan el método de la ciencia como *la* manera que nos permite: *a)* plantear cuestiones fácticas "razonables" (esto es, preguntas que son significativas, no triviales, y que probablemente pueden ser respondidas dentro de una teoría existente o concebible); y *b)* probar respuestas probables en todos los campos especiales del conocimiento.

No debe confundirse la filosofía científica con el *cientificismo ingenuo* en cualquiera de sus dos versiones: el enciclopedismo científico y el reduccionismo naturalista. El enciclopedismo científico pretende que la única tarea de los filósofos es recoger los resultados más generales de la ciencia, elaborando una imagen unificada de los mismos, y preferiblemente formulándolos todos en un único lenguaje (p. ej., el de la física). En cambio, la filosofía, científica o no, analiza lo que se le presenta y, a partir de este material, construye teorías de segundo nivel, es decir, teorías de teorías; la filosofía será científica en la medida en que elabore de manera racional los materiales previamente elaborados por la ciencia. Así es como puede entenderse la extensión del método científico al trabajo filosófico.

En cuanto al cientificismo concebido como reduccionismo naturalista —y que a veces se superpone con el enciclopedismo científico, como ocurre con el fisicismo—, puede describírselo como una tentativa de resolver toda suerte de problemas con ayuda de las técnicas creadas por las ciencias naturales, desdeñando las cualidades específicas, irreductibles, de cada nivel de la realidad. El cientificismo radical de esta especie sostendría, por ejemplo, que la sociedad no es más que un sistema físico químico (o, a lo sumo, biológico), de donde los fenómenos sociales debieran estudiarse exclusivamente mediante la ayuda de metros, relojes, balanzas y otros instrumentos de la misma clase. En cambio, la filosofía científica favorece la elaboración de técnicas específicas en cada campo, con

la única condición de que estas técnicas cumplan las exigencias esenciales del método científico en lo que respecta a las preguntas y a las pruebas. De esta manera es como puede entenderse la extensión del método científico a todos los campos especiales del conocimiento.

Pero también debiera emplearse el método de la ciencia en las ciencias aplicadas y, en general, en toda empresa humana en que la razón haya de casarse con la experiencia; vale decir, en todos los campos excepto en arte, religión y amor. Una adquisición reciente del método científico es la investigación operativa *(operations research)*, esto es, el conjunto de procedimientos mediante los cuales los dirigentes de empresas pueden obtener un fundamento cuantitativo para tomar decisiones y los administradores pueden adquirir ideas para mejorar la eficiencia de la organización.[10] Pero, desde luego, la extensión del método científico a las cosas humanas está aún en su infancia. Pídasele a un político que pruebe sus afirmaciones, no recurriendo a citas y discursos, sino confrontándolas con hechos certificables (tal como se recogen y elaboran, p. ej., con ayuda de las técnicas estadísticas). Si es honesto, cosa que puede suceder, o bien *a)* admitirá que no entiende la pregunta, o *b)* concederá que todas sus creencias son, en el mejor de los casos, enunciados plausibles, ya que sólo pue-

[10] Véase P. M. Morse y G. E. Kimball, *Methods of Operations Research*, ed. rev. (Cambridge, Mass., The Technology Press of Massachusetts Institute of Technology; Nueva York, John Wiley & Sons, 1951).

den ser probados imperfectamente, o *c)* llegará a la conclusión de que muchas de sus hipótesis favoritas (principios, máximas, consignas) tienen necesidad urgente de reparación. En este último caso puede terminar por admitir que una de las virtudes del método de la ciencia es que facilita la regulación o readaptación de las ideas generales que guían (o justifican) nuestra conducta consciente, de manera tal que ésta pueda corregirse con el fin de mejorar los resultados.

Desgraciadamente, la cientifización de la política la haría más eficaz, pero no necesariamente mejor, porque el método puede dar la forma y no el contenido; y el contenido de la política está determinado por intereses que no son primordialmente culturales o éticos, sino materiales. Por esto, una política científica puede dirigirse a favor o en contra de cualquier grupo social: los objetivos de la estrategia política, así como los de la investigación científica aplicada, no son fijados por patrones científicos, sino por intereses sociales. Esto muestra a la vez el alcance y los límites del método científico: por una parte, puede producir saber, eficiencia y poder; por la otra, este saber, esta eficiencia y este poder pueden usarse para bien o para mal, para libertar o para esclavizar.

12. El método científico: ¿un dogma más?

¿Es dogmático favorecer la extensión del método científico a todos los campos del pensamiento

y de la acción consciente? Planteamos la cuestión en términos de conducta. El dogmático vuelve sempiternamente a sus escrituras, sagradas o profanas, en búsqueda de la verdad; la realidad le quemaría los papeles en los que imagina que está enterrada la verdad: por esto elude el contacto con los hechos. En cambio, para el partidario de la filosofía científica todo es problemático: todo conocimiento fáctico es falible (pero perfectible), y aun las estructuras formales pueden reagruparse de maneras más económicas y racionales; más aún, el propio método de la ciencia será considerado por él como perfectible, como lo muestra la reciente incorporación de conceptos y técnicas estadísticas. Por consiguiente, el partidario del método científico no se apegará obstinadamente al saber, ni siquiera a los medios consagrados para adquirir conocimiento, sino que adoptará una actitud investigadora; se esforzará por aumentar y renovar sus contactos con los hechos y el almacén de las ideas mediante las cuales los hechos pueden entenderse, controlarse y a veces reproducirse.

No se conoce otro remedio eficaz contra la fosilización del dogma —religioso, político, filosófico o científico— que el método científico, porque es el único procedimiento que no pretende dar resultados definitivos. El creyente busca la paz en la aquiescencia; el investigador, en cambio, no encuentra paz fuera de la investigación y de la disensión: está en continuo conflicto consigo mismo, puesto que la exigencia de buscar conocimiento verificable implica un continuo inventar, probar y

criticar hipótesis. Afirmar y asentir es más fácil que probar y disentir; por esto hay más creyentes que sabios, y por esto, aunque el método científico es opuesto al dogma, ningún científico y ningún filósofo científico debieran tener la plena seguridad de que han evitado todo dogma.

De acuerdo con la filosofía científica, el peso de los enunciados —y por consiguiente su credibilidad y su eventual eficacia práctica— depende de su grado de sustentación y de confirmación. Si, como estimaba Demócrito, una sola demostración vale más que el reino de los persas, puede estimarse el valor del método científico en los tiempos modernos. Quienes lo ignoran íntegramente no pueden llamarse modernos; y quienes lo desdeñan se exponen a no ser veraces ni eficaces.

5.1.2. Formulación de preguntas de investigación derivadas del problema identificado

Para el desarrollo del subtema **Formulación de preguntas de investigación derivadas del problema identificado**, se consideran los contenidos relacionados con el **planteamiento y formulación del problema**, así como con la **delimitación del problema o de la investigación**, debido a que estos permiten partir de una situación problemática previamente identificada, precisarla, establecer sus límites y convertirla en una base adecuada para la formulación de las preguntas de investigación.

De acuerdo con las obras de **Fidias G. Arias** y **César Augusto Bernal Torres**, los contenidos relacionados con este subtema se encuentran en los capítulos y apartados que se presentan a continuación:

N.°	Libro	Capítulo	Página	Contenido
1	*El Proyecto de Investigación. Introducción a la metodología científica* — **Fidias G. Arias**	**Capítulo 3: El problema de investigación**	**p. 41**	**3.2. Planteamiento y formulación del problema**
2	*El Proyecto de Investigación. Introducción a la metodología científica* — **Fidias G. Arias**	**Capítulo 3: El problema de investigación**	**p. 42**	**3.3. Delimitación del problema**
3	*Metodología de la investigación* — **César Augusto Bernal Torres**	**Capítulo 7: Proceso de investigación científica**	**p. 88**	**7.2. Problema de la investigación**
4	*Metodología de la investigación* — **César Augusto Bernal Torres**	**Capítulo 7: Proceso de investigación científica**	**p. 106**	**7.4. Justificación y delimitación de la investigación — solo delimitación de la investigación**

Autores considerados

- **Autor principal:** Fidias G. Arias
  Obra: *El Proyecto de Investigación. Introducción a la metodología científica*, 6.ª edición.
- **Autor de apoyo:** César Augusto Bernal Torres
  Obra: *Metodología de la investigación*.

Ejemplo:

Preguntas formuladas de manera incorrecta:

a) ¿Contribuye la lectura al desarrollo de la memoria?

b) ¿Influye el entorno familiar en el rendimiento escolar?

Preguntas reformuladas:

a) ¿Cuál es la relación entre lectura y desarrollo de la memoria?

b) ¿Cómo influye el entorno familiar en el rendimiento escolar?

8. La pregunta debe estar delimitada, es decir, incluirá con precisión el espacio, la población y el tiempo o período al que se refiere (ver sección 3.3).

3.2. Planteamiento y formulación del problema

Aunque para algunos autores, planteamiento y formulación del problema son términos equivalentes, en esta obra se consideró pertinente diferenciarlos.

> El planteamiento del problema consiste en describir de manera amplia la situación objeto de estudio, ubicándola en un contexto que permita comprender su origen, relaciones e incógnitas por responder.

Plantear el problema implica desarrollar, explicar o exponer con amplitud. Mientras que formular es concretar, precisar o enunciar. En este sentido:

> Formulación del problema es la concreción del planteamiento en una pregunta precisa y delimitada en cuanto a espacio, tiempo y población (si fuere el caso).

Puede ocurrir que la formulación contenga más de una pregunta. Lo indispensable es que exista una estrecha relación entre las interrogantes formuladas.

3.3. Delimitación del problema

Al igual que el tema, el problema también debe ser delimitado. En el campo de la investigación, delimitar implica establecer los alcances y límites en cuanto a lo que se pretende abarcar en el estudio. Concretamente:

> La delimitación del problema significa indicar con precisión en la interrogante formulada: el espacio, el tiempo o período que será considerado en la investigación, y la población involucrada (si fuere el caso).

3.3.1. Delimitación del espacio

Es muy importante que la pregunta precise el ámbito o lugar que será tratado en el estudio.

Ejemplo:

¿Cuál fue la tasa de analfabetismo en el Estado Vargas - Venezuela para el año 2003?

3.3.2. Delimitación de tiempo

En la formulación del problema debe indicarse el lapso o período objeto de estudio.

Ejemplo:

¿Cuáles fueron las causas del alto número de aplazados en el "Instituto Universitario Académico" durante el semestre octubre 2011- marzo 2012?

3.3.3. Delimitación de la población

En este caso hay que señalar los sujetos que serán observados, encuestados o medidos:

Ejemplo:

¿Cuál es el nivel de aptitud física de los alumnos del 1er semestre del Colegio Universitario de Caracas?

Es importante señalar que algunos problemas no incluyen los tres elementos de una delimitación. El objeto de estudio determinará la pertinencia de incluirlos o no. Ejemplo:

¿Qué factores ocasionaron las pérdidas en la "Empresa HL" durante el período diciembre 2011 - febrero 2012?

En este caso se precisa el espacio: "Empresa HL", y el tiempo: período diciembre 2011 - febrero 2012. Sin embargo, no se hace referencia a una población específica, lo que también resulta válido.

3.4. Objetivos de investigación

Objetivo es sinónimo de meta, es decir, aquello que se aspira lograr o alcanzar. En este caso, nos referimos a meta en términos de conocimiento, es decir, los conocimientos que el investigador pretende obtener. En este orden de ideas:

> Objetivo de investigación es un enunciado que expresa lo que se desea indagar y conocer para responder a un problema planteado.

3.4.1. Características de los objetivos de investigación

a) Indican los conceptos que serán estudiados.
b) Precisan las variables o dimensiones que serán medidas.
c) Señalan los resultados que se esperan.
d) Definen los límites o alcances de la investigación.
d) Se redactan comenzando con un verbo en infinitivo.
e) Deben ser posibles de lograr.
f) Junto al problema de investigación, los objetivos responden a la pregunta ¿qué se pretende con la investigación? (Hernández y otros 2010; Ramírez, 2010; Sabino 2006), y no al ¿para qué?, como señalan algunos autores (ver ejemplo en la página 97).

A continuación se presenta una lista de verbos indicados para objetivos de investigación, clasificados según el nivel (Arias, 2006 b).

Nivel Exploratorio	Nivel Descriptivo	Nivel Explicativo
Conocer	Analizar	Comprobar
Definir	Calcular	Demostrar
Descubrir	Caracterizar	Determinar
Detectar	Clasificar	Establecer
Estudiar	Comparar	Evaluar
Explorar	Cuantificar	Explicar
Indagar	Describir	Inferir
Sondear	Diagnosticar	Relacionar
	Examinar	Verificar
	Identificar	
	Medir*	

Debe evitarse confundir los objetivos de investigación con:

a) Objetivos educativos o instruccionales: expresan lo que el alumno debe lograr como producto del proceso de enseñanza-aprendizaje.

Ejemplo:

"Al finalizar esta unidad, el alumno estará en capacidad de redactar oraciones completas..."

b) Objetivos prácticos: implican una habilidad psicomotora y son típicos de los proyectos tecnológicos que se desarrollan en áreas como Ingeniería, Electrónica y Computación.

Ejemplos:

"Diseñar un prototipo..."

"Desarrollar un programa computarizado (software)..."

*El verbo medir no debe ser entendido como la simple actividad que implica utilizar una cinta métrica para obtener la estatura de una persona, o las dimensiones de un terreno. En investigación social, medir significa la ejecución de objetivos de gran complejidad. Por ejemplo: medir la pobreza; medir la eficiencia, entre otros.

7.2 Problema de la investigación

Para que una idea sea objeto de investigación, debe convertirse en ***problema de investigación***. Ahora, en investigación, problema es todo aquello que se convierte en objeto de reflexión y sobre el cual se percibe la necesidad de conocer y, por tanto, de estudiar.

En este sentido, problema no es algo disfuncional, molesto o negativo, sino todo aquello que incite a ser conocido, pero teniendo en cuenta que su solución sea útil, es decir, buscando una respuesta que resuelva algo práctico o teórico. Por esto, a este modelo de investigación, además de ser conocido como modelo general, también suele denominársele *modelo pragmático*.

Una vez definidos el tema y el título de la *propuesta de investigación* se procede a plantear el problema de investigación, entendiendo como problema de investigación la situación, el fenómeno, el evento, el hecho u objeto del estudio que se va a realizar.

Arias Galicia (1991) considera que "en investigación no es suficiente visualizar un problema, es necesario plantearlo adecuadamente" (p. 50). Los especialistas en investigación enfatizan en la necesidad de un buen planteamiento del problema; para ellos, si esto se logra, la mitad del problema se ha solucionado. En este sentido, Briones (1985) afirma que "el planteamiento de un problema es la fase más importante de todo el proceso de investigación" (p. 39).

Plantear el problema de investigación significa enunciar y formular el problema, aspectos que se definen a continuación.

7.2.1 Enunciar el problema

Enunciar un problema de investigación consiste en presentar, mostrar y exponer las características o los rasgos del tema, situación o aspecto de interés que va a estudiarse, es decir, describir el estado actual del problema.

En general, enunciar un problema es contar lo que está pasando en relación con una situación, con una persona o con una institución; es narrar los hechos que caracterizan esa situación, mostrando sus implicaciones y soluciones.

Enunciar un problema requiere precisar la naturaleza y las dimensiones del mismo, en detalle y con precisión. Asimismo, se deben ambientar todas las características que enmarcan el problema; también hay que comenzar por narrar los antecedentes de la situación de estudio, así como incluir y mostrar los hechos, las relaciones y las explicaciones que sean importantes en la caracterización del problema. Igualmente, hay que contemplar tanto el problema como los elementos conectados con él (Tamayo, 2002).

www.FreeLibros.me

Tamayo (2002), citando a Van Dalen , sugiere tener en cuenta los siguientes aspectos al momento de plantear o definir un problema de investigación:

1. Reunir los hechos en relación con el problema (qué está pasando).
2. Determinar la importancia de los hechos.
3. Identificar las posibles relaciones entre los hechos que pueden indicar la causa de la dificultad.
4. Proponer explicaciones para conocer la causa de la dificultad y determinar su importancia en el problema.
5. Encontrar, entre las explicaciones, aquellas relaciones que permitan adquirir una visión amplia de la solución del problema.
6. Hallar relaciones entre los hechos y las explicaciones.
7. Analizar los supuestos en los que se apoyan los elementos identificados.

Para lograr definir o describir bien el problema hay que poseer un conocimiento previo sobre la situación que se va a estudiar, lo cual habrá de reflejarse en un conocimiento general que debe tener el investigador sobre el objeto o sujeto que se va a estudiar y sobre investigaciones específicas adelantadas en el asunto, así como algunas experiencias personales.

En síntesis, enunciar un problema es presentar una *descripción* general de la situación objeto de estudio.

7.2.2 Formular el problema

Un problema se *formula* cuando el investigador dictamina o hace una especie de *pronóstico* sobre la situación problema. En lugar de hacerlo con afirmaciones, este pronóstico se plantea mediante la formulación de preguntas orientadas a dar respuesta al problema de la investigación (Méndez, 1995).

Una adecuada formulación de un problema de investigación implica elaborar dos niveles de preguntas. La pregunta general debe recoger la esencia del problema y, por tanto, el título del estudio.

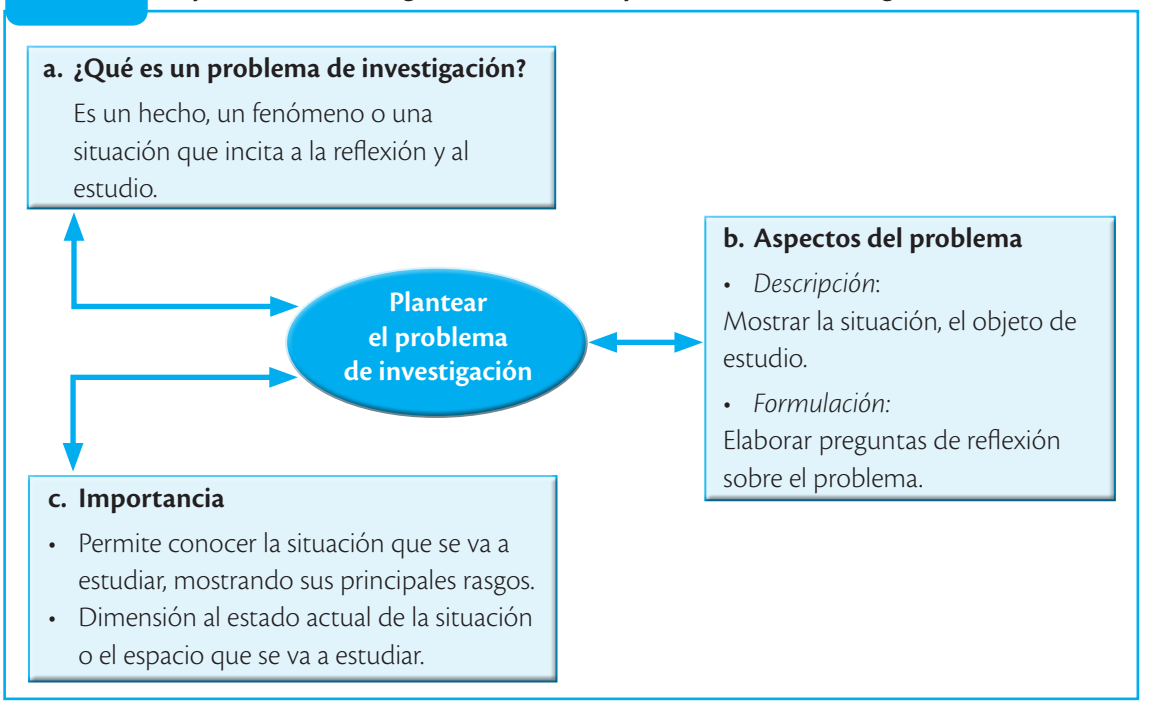
Las preguntas específicas están orientadas a interrogar sobre aspectos concretos del problema y no al problema en su totalidad, pero que en su conjunto conforman la totalidad (las preguntas específicas son subpreguntas de la pregunta general).

La Figura 7.4 ilustra los aspectos relevantes relacionados con el planteamiento del problema de investigación tales como el propio concepto de problema en investigación, los aspectos del problema y la importancia que representa una adecuada definición del problema en un proceso de investigación científica.

En los ejemplos 7.3 y 7.4 se muestran planteamientos de problemas de investigación en un proyecto de investigación:

www.FreeLibros.me

FIGURA 7.4 **El proceso de investigación. Plantear el problema de la investigación**



EJEMPLO 7.3 **Planteamiento del problema de investigación en el campo de la psicología y la educación**

Título de la investigación: *Compromiso pedagógico de la literatura infantil latinoamericana producida en la década de 1990 en la formación integral del niño y la niña.*

Planteamiento del problema

Enunciado del problema

Los niños viven en nuestros días en medio de una civilización tecnológica o postindustrial, rodeada cotidianamente por un flujo incesante y multiforme de estímulos y mensajes, así como inmersa en un perturbador y seductor universo de imágenes. Los medios de comunicación, en sus diversas expresiones y articulaciones, invaden, como protagonistas, cada aspecto y momento de la vida de las personas, en especial de la población infantil.

No extraña que en los países industrializados, como afirma Postman (1998), la formación del niño se deba en un 75% a la influencia de los medios de comunicación y sólo en un 25% a las instancias educativas tradicionales, mientras que a comienzos de la década de 1950 las proporciones eran inversas.

www.FreeLibros.me

Con la irrupción de la galaxia informática, telematizada, se anuncia, cargada de incógnitas y no carente de riesgos, una posterior revolución de la comunicación antropológica cultural, lógica-cognitiva y sensorioperceptiva, de proporciones inadvertidas (Nobile 1992) .

La presencia invasora de los medios de comunicación, complicada por la revolución telemática e informativa, operando junto a otros factores y diversas situaciones, propios de la actual civilización tecnológica, produce consecuencias claras en el desarrollo psicológico de la infancia y en su condición existencial. En el plano intelectual, hay que señalar la pasividad de recepción, el empobrecimiento del espíritu crítico y la inhibición de la autonomía de juicio, los cuales, unidos al empobrecimiento de la fantasía y la anulación de la creatividad, se traducen en un obstáculo para el desarrollo, la integración y el perfeccionamiento de las facultades superiores de la inteligencia y del pensamiento.

En el plano estrictamente intelectual, según diversos autores, la excesiva familiaridad del infante con la imagen no favorece el desarrollo cognitivo ni el de la fantasía; incluso, se perjudica el pensamiento creativo, determinando, a la vez, el progresivo deterioro de la curiosidad infantil.

La aparente fluidez expresiva lingüística, la riqueza léxica, el uso de un lenguaje más elaborado del niño televisivo no implican estructuras cognitivas más complejas, y son el resultado de adquisiciones superficiales y efímeras.

En tal contexto, una ponderada reflexión sobre la problemática sociopsicopedagógica de la civilización postindustrial, en especial en un deteriorado aspecto de la sociedad de consumo, hace recuperar y destacar el papel de la lectura crítica e inteligente, en la acepción bruneriana del término, en función antagónica y correctora de los nefastos efectos producidos por la invasora difusión de los medios audiovisuales, así como de la anunciada hegemonía de los nuevos medios electrónicos omnipresentes y multiperceptivos.

Si el libro es "un medio personal, capaz de poner en situación de objetiva autonomía cultural al usuario", la lectura, coloquio íntimo con el autor, reivindica su permanente función exquisitamente formativa como factor esencial de autorrealización de la persona, "actitud que implica toda la vida psíquica" (Guini, 1969), rica en reclamos interiores e interiorizados, principal instrumento de enculturización, proceso de reelaboración, interpretación y sistematización de la realidad, aproximación viva y vivificante a la lengua, que asume la irrenunciable función de ampliación de los conocimientos, así como de los horizontes intelectuales y culturales, de estructuración y potenciación de las facultades lógicas, de enriquecimiento del patrimonio lingüístico y de los medios expresivos.

Al ejercitarse en textos cualificados, en cuanto a lenguaje y contenidos, la lectura agudiza el espíritu crítico, refuerza la autonomía de juicio, educa el sentimiento estético, nutre la fantasía, ensancha la imaginación, habla a la afectividad, cultiva el sentimiento, descubre intereses más amplios y autónomos, contribuye a la promoción de una sólida conciencia moral y cívica, abierta a los ideales de comprensión humana y de solidaridad, por lo que resulta ideal para la formación integral de la persona, en especial cuando está en edad de formación.

www.FreeLibros.me

Al mismo tiempo, en el clamor ensordecedor de los medios de comunicación, la lectura permite el redescubrimiento de los valores psicopedagógicos del silencio y, algo aún más urgente, la recuperación de la dimensión de interioridad, favoreciendo el hábito de la reflexión y la introspección.

En una edad evolutiva, marcada por un flujo incesante de imágenes y acontecimientos, así como sumergida en un desestructurado y desestructurante torrente de información, que acaba por configurar todo de una forma confusa y fragmentaria, irreflexiva y pasiva, la lectura de calidad aporta los instrumentos, especialmente lógico-formales y lingüísticos, para su asimilación crítica e informada, lo mismo que para un uso activo de esos mismos medios, apareciendo como elemento de orden y disciplina interior.

En el siglo XXI , como supone Frabioni (1987), ante el niño y la niña, sentados en solitario delante de sus propios visos, pasarán miles de páginas de información memorizadas en gigantescos bancos de datos. Por otro lado, frente a un descenso general de las relaciones interpersonales, de convivencia y de posibilidades sociales de juegos de grupo se prevé el empobrecimiento y la fragmentación del saber, una invasión de imágenes y palabras escritas preelaboradas, con el consiguiente descenso de la creatividad y el desarrollo integral de la persona.

Las llamadas tecnologías de punta, especialmente la microelectrónica y la microinformática, tienden a influir y a acondicionar formas de razonamiento y pensamiento, dando mayor importancia al sistema informático y una comprensión unidimensional de la realidad.

El libro, depositario de un inestable patrimonio de conocimientos, cultura y civilización, lo mismo que síntesis de las creaciones de la fantasía y el ingenio humanos, es un medio fundamental privilegiado para cumplir una labor universal de educación, de entenderse como una transfusión de vida de una generación a otra.

Así como en la narrativa serena y optimista, lo mismo que rica en valores humanos, sociales y cívicos, además de interesante y atractiva, que plantea, a través de conocimientos culturales elegidos adecuadamente, ideales de vida positivos y modelos de conducta e identificación, indispensables en el proceso evolutivo, que ayuden en forma gradual al joven lector a tener contacto con la realidad y sus problemas, pueden ser un elemento importante en el desarrollo armónico e integral de la personalidad infantil, y asumir la urgente tarea de la regeneración espiritual de la juventud.

La literatura, especialmente la infantil, afirma Nobile (1992), debe alzarse en una sólida unidad de intenciones con la escuela, como último y generoso baluarte, contra la creciente marca del instinto, la ignorancia, la simultánea y destructora amenaza de un empobrecimiento humano y, sobre todo, el grave deterioro de la palabra como hecho espiritual, como descubrimiento personal y como fuente de desarrollo humano.

Cuando se habla aquí de literatura, se hace referencia a una literatura auténticamente formativa, capaz de ofrecer valores y contenidos culturales a una población infantil con un perfil cada vez más conformista, estandarizado y homogenizado, tanto en el lenguaje como en el vestido y en los comportamientos, lo mismo que llena de mitos consumistas por la persistencia del

www.FreeLibros.me

originario principio del placer, atento sólo a lo inmediato y egoísta, satisfacción de las propias e indisciplinadas pulsaciones del instinto.

En este sentido, la literatura es y debe ser fuente de expresión de humanidad y de racionalidad crítica, garantía de libertad de pensamiento, así como fuente inagotable de desarrollo humano integral; por ello y las razones expresadas que hoy enfrenta la población infantil, es importante estudiar el compromiso pedagógico de la literatura infantil en la formación de sus lectores.

Formulación del problema

En virtud de lo anterior, el presente trabajo se orienta al análisis formativo de la literatura infantil (latinoamericana), pero específicamente a la promoción de valores en los niños y niñas; por tanto, el estudio busca responder los siguientes interrogantes:

- ¿Qué se entiende por literatura infantil?
- ¿Existe compromiso pedagógico de la literatura infantil latinoamericana de la década de 1990 en la formación integral de sus lectores?
- ¿Hay relación entre la literatura infantil latinoamericana de la década del 90 y los valores que practican los niños y niñas que leen esta literatura?
- ¿La literatura infantil latinoamericana de la década del 90 ha sido un instrumento pedagógico de promoción de la dimensión social y psicoafectiva para el niño o la niña que lee esta literatura?

EJEMPLO 7.4 Planteamiento del problema de investigación aplicado al campo de los negocios

En el caso de una investigación sobre creatividad en las empresas nacionales, cuyo título podría ser *La industria manufacturera nacional frente a la innovación y la creatividad*, el planteamiento del problema sería el siguiente:

Planteamiento del problema

Enunciado del problema

En un mundo caracterizado por rápidos y complejos cambios, a los directivos empresariales cada vez se les plantean problemas más difíciles que no sólo afectan su organización, sino que influyen en el conjunto de la sociedad.

Por esto, cada vez tendrán mayor necesidad de enfrentar las situaciones de un modo creativo e innovador. Sin embargo, estudios recientes sobre el manejo de la creatividad y la

www.FreeLibros.me

innovación en las organizaciones empresariales revelan que en realidad el camino por recorrer es muy largo.

Investigadores como Sternberg y Lubar, y Kuczmarski (1997), en sus estudios sobre creatividad en las organizaciones, concluyen que la mayoría de los gerentes de alto nivel se atemorizan ante las personas creativas e innovadoras, pues las consideran incómodas y raras, y se niegan a abrazar lo novedoso y diferente, cuando desaniman y no premian el pensamiento innovador de sus empleados.

En general, pese a no reconocerlo, los directivos de las empresas nacionales entorpecen o no propician la creatividad, lo cual los pone en desventaja para asumir los retos del nuevo milenio.

La sociedad actual necesita directivos capaces de asumir un verdadero compromiso con sus trabajadores, más que desempeñar funciones preestablecidas, que sean innovadores y comprometidos con su propio desarrollo y el de la sociedad.

Formulación del problema

- ¿Tienen los empresarios nacionales de la industria manufacturera una actitud congruente entre lo que piensan y lo que hacen respecto a las personas creativas en sus empresas? (Formulación de una pregunta de investigación general).
- ¿Tienen los empresarios nacionales actitudes positivas frente a las personas creativas? (Pregunta específica).
- ¿Estimulan y promueven los empresarios nacionales la creatividad? (Pregunta específica).

Otro planteamiento de un problema de investigación podría ser el que se describe a continuación y que tiene que ver con los *Nuevos retos que enfrentan las empresas como consecuencia del nuevo paradigma mundial de la globalización y la competitividad.*

Planteamiento del problema

Enunciado del problema

Reconocidos investigadores en el ámbito de los negocios y de la economía, como Kenich Ohmae, Peter F. Drucker, Patricia Aburdene, John Naisbitt, Jack Nadel y Michael E. Porter, han analizado los cambios económicos del mundo, la naturaleza de los negocios y la necesidad de tener un nuevo paradigma, un nuevo enfoque y una nueva manera de pensar para interpretar y abordar la nueva realidad.

El planeta Tierra parece haberse convertido en una pequeña aldea. El avance de la globalización es una realidad; los individuos ya no son ciudadanos de una nación, sino ciudadanos del mundo; los sistemas de comunicación vía satélite, la telefonía celular, las autopistas de información vía Internet, la fibra óptica y los computadores están revolucionando los procesos de conocimiento.

www.FreeLibros.me

En el mundo de la política están rompiéndose los paradigmas básicos. Las estructuras de poder ligadas a grandes organizaciones sociales como partidos políticos y sindicatos dan paso ahora a los candidatos libres. Ya la fuerza no está en los partidos políticos, ahora está en las personas, en sus características humanas, individuales, personales.

Cambian los estilos de vida, se da una transformación total de los sistemas educativos del mundo. La calidad de la educación se caracterizará por un sistema de competencias individuales y por su infraestructura orientada hacia la investigación.

Se está en los albores de la era de la creatividad, de la era del conocimiento. Los países que tengan los mejores científicos, universidades, centros de investigación y centros de alta tecnología dominarán el mundo.

En el siglo XXI se dará un salto cualitativo en cuanto al aprovechamiento de las energías eólica y solar, y habrá mayor posibilidad de integración armónica con el entorno, y con formas energéticas menos agresivas para el medio biológico y para la biosfera.

Los negocios están enfrentando una paradoja, pues tienen oportunidades nunca vistas para aprovechar los nuevos mercados; sin embargo, éstos cambian de manera sustancial, reduciéndose o haciéndose intensamente competitivos. La apertura de los mercados mundiales generará una reestructuración en cada sector de los negocios, orientándolos a garantizar las siguientes estrategias: rapidez, bajo margen de ganancia por unidad, calidad en productos y servicios, diversidad de productos, descuentos, comodidad e innovación (Tucker, 1995).

Formulación del problema

- Frente a esta nueva realidad, ¿cómo deben prepararse las empresas nacionales y sus gerentes si quieren desenvolverse competitivamente? (Pregunta general).
- ¿Cómo favorece el gobierno la competitividad de las empresas nacionales?
- ¿Cuál es la actitud de los gremios ante los retos que impone el nuevo milenio?
- ¿Qué están haciendo las universidades para contribuir a la competitividad de las empresas?
- ¿Qué están haciendo los empresarios para responder al nuevo ambiente de competitividad?

RESUMEN

En investigación, problema es todo aquello que se convierte en objeto de reflexión y sobre el cual se percibe la necesidad de conocer y, por tanto, de estudiar. Es decir, problema no significa algo disfuncional, molesto o negativo, sino todo aquello que incite a ser conocido, y su conocimiento representa alguna utilidad ya sea para el investigador o para un tercero. De otra parte, en investigación no es suficiente visualizar un problema, es necesario plantearlo adecuadamente.

www.FreeLibros.me

7.4 Justificación y delimitación de la investigación

Toda investigación está orientada a la resolución de algún problema; por consiguiente, es necesario *justificar,* o exponer, los motivos que merecen la investigación. Asimismo, debe determinarse su cubrimiento o dimensión para conocer su viabilidad.

7.4.1 Criterios de justificación

De acuerdo con Méndez, la justificación de una investigación puede ser de carácter teórico, práctico o metodológico (Méndez, 1995).

Justificación teórica

En investigación hay una *justificación teórica* cuando el propósito del estudio es generar reflexión y debate académico sobre el conocimiento existente, confrontar una teoría, contrastar resultados o hacer epistemología del conocimiento existente.

En el caso de las ciencias económico-administrativas, un trabajo investigativo tiene justificación teórica cuando se cuestiona una teoría administrativa o una económica (es decir, los principios que la soportan), su proceso de implantación o sus resultados.

Cuando en una investigación se busca mostrar las soluciones de un modelo, está haciéndose una justificación teórica.

Si en una investigación se proponen nuevos paradigmas o se hace una reflexión epistemológica, se tiene una justificación eminentemente teórica, aunque al implementarla se vuelve práctica, ya que, como afirma López Cerezo (1988), toda investigación en alguna medida tiene la doble implicación, teórica y práctica.

La justificación teórica es la base de los programas de doctorado y algunos programas de maestría donde se tiene como objetivo la reflexión académica.

Justificación práctica

Se considera que una investigación tiene *justificación práctica* cuando su desarrollo ayuda a resolver un problema o, por lo menos, propone estrategias que al aplicarse contribuirían a resolverlo.

Los estudios de investigación de pregrado y de posgrado, en el campo de las ciencias económicas y administrativas, en general son de carácter práctico, o bien, describen o analizan un problema o plantean estrategias que podrían solucionar problemas reales si se llevaran a cabo.

Cuando en un trabajo de grado se realiza un análisis económico de un sector de la producción, su justificación es práctica porque genera información que podría utilizarse para tomar medidas tendientes a mejorar ese sector.

www.FreeLibros.me

Cuando un trabajo de grado se orienta a conocer los factores de motivación más utilizados en un determinado sector económico o en una empresa, su justificación es práctica, porque, al igual que en el caso del análisis del sector, la información sirve para actuar sobre la empresa, para mejorar o realizar *benchmarking* en otras organizaciones, o para confrontar la teoría sobre el tema.

Justificación metodológica

En investigación científica, la *justificación metodológica* del estudio se da cuando el proyecto que se va a realizar propone un nuevo método o una nueva estrategia para generar conocimiento válido y confiable.

7.4.2 Limitaciones del estudio o de la investigación

Una vez justificada la investigación, es necesario plantear las limitaciones dentro de las cuales se realizará (no todos los estudios tienen las mismas limitaciones, pues cada investigación es particular).

Las limitaciones en un proyecto de investigación pueden referirse a:

Limitaciones de tiempo Es necesario determinar cuál será el período, sea retrospectivo o prospectivo, dentro del cual se realizará el estudio del hecho, la situación, el fenómeno o población investigados. Por ejemplo, si el interés es estudiar el comportamiento de un sector económico, o los indicadores económicos de un país, es indispensable definir durante qué período (años, meses) se realizará tal análisis.

Limitaciones de espacio o territorio Son aquellas demarcaciones del espacio geográfico dentro del cual tendrá lugar una investigación. Las investigaciones pueden limitarse a una zona de una ciudad, a una ciudad, una región, un país, un continente, etcétera.

Limitaciones de recursos Mencionan la disponibilidad de recursos financieros para la realización del proyecto de investigación.

Además de estas limitaciones, en un proyecto de investigación puede haber limitaciones de información, población disponible para el estudio, dificultad de acceso a la misma, etcétera.

El investigador debe explicitar las limitaciones del proyecto con el propósito de facilitar su viabilidad.

EJEMPLO 7.9 **Enunciados de limitaciones de tiempo**

- Diagnóstico del sector de las flores en los últimos 12 años (1997-2009).
- Evaluación de la tasa de desempleo durante los últimos cuatro gobiernos.
- Análisis de las estrategias didácticas utilizadas en el proceso enseñanza-aprendizaje de la educación superior en los últimos 10 años.

www.FreeLibros.me

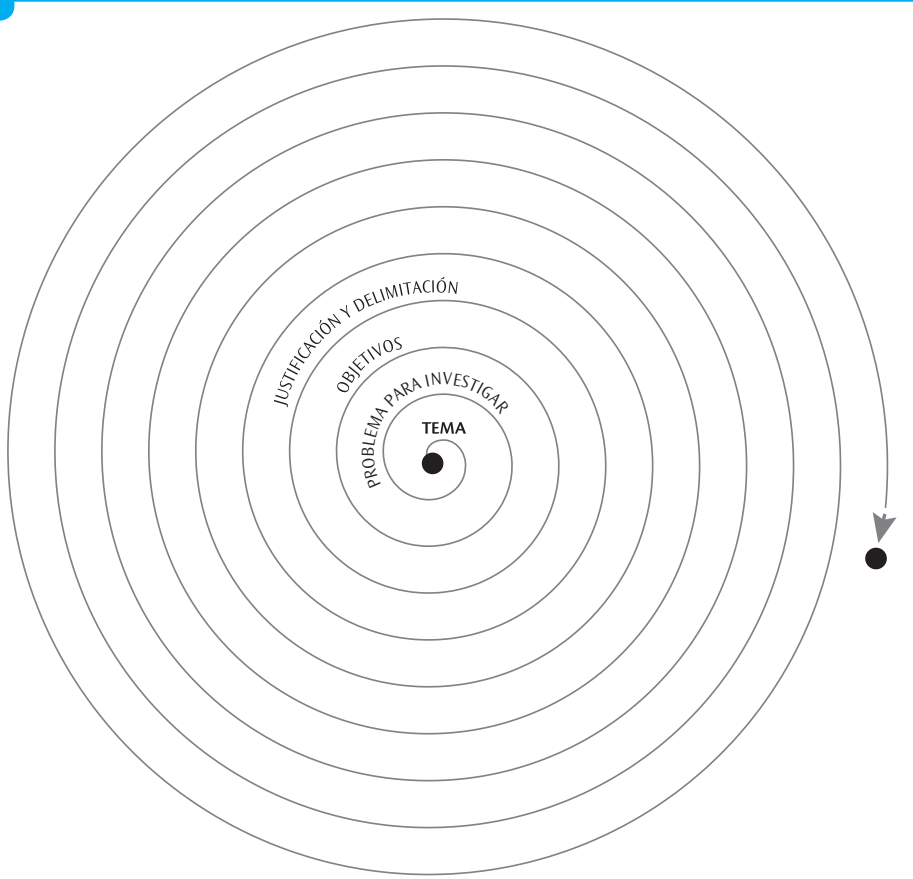
Existen investigaciones que se circunscriben a una determinada región, y eso se conoce como limitaciones geográficas.

EJEMPLO 7.10 **Enunciados de limitaciones de espacio o territorio**

- Cuando se dice que el estudio se realizará en empresas localizadas en la ciudad *PQTL* o en las empresas latinoamericanas, se hace referencia a una limitación geográfica.
- Afirmar que un estudio se realizará en una determinada ciudad o zona geográfica implica delimitar este estudio geográfica o territorialmente.

En general, puede hablarse de varias limitaciones según las características de la población o muestra sobre la cual recae el estudio. Estas características se detallarán cuando se haga referencia a la población y muestra objeto de estudio.

FIGURA 7.7 **El proceso de investigación. Justificación y delimitación de la investigación**



www.FreeLibros.me

La Figura 7.7 sintetiza la justificación como las razones de carácter práctico, teórico y metodológico que se tienen para realizar el estudio, así como sus delimitaciones en los ámbitos espacial, cronológico y/o sociodemográfico.

RESUMEN

En investigación, la justificación se refiere a las razones del porqué y el para qué de la investigación que se va a realizar, es decir, justificar una investigación consiste en exponer los motivos por los cuales es importante llevar a cabo el respectivo estudio. Al respecto, suele haber tres dimensiones o tipos de justificación: teórica, práctica y metodológica. Algunas investigaciones pueden requerir los tres tipos de justificación, otras dos tipos y otras solo uno. Ello depende de las particularidades de cada investigación.

La delimitación o el alcance en investigación se refiere a la dimensión o al cubrimiento que ésta tendrá en el espacio geográfico, período de tiempo y perfil sociodemográfico del objeto de estudio.

EJERCICIOS DE REPASO Y ANÁLISIS

1. ¿Qué significa justificar y delimitar una investigación?
2. Explique las distintas formas de justificar una investigación.
3. Seleccione un tema de investigación y justifique por qué debe realizarse el estudio.
4. Revise los ejemplos de los temas propuestos en el tema relacionado con la formulación de objetivos que justifiquen la realización de esos estudios.
5. Suponga que quiere realizar un estudio sobre el concepto de ser humano que han manejado las distintas teorías administrativas a través del tiempo. ¿Cuál es la justificación más adecuada para este estudio? Justifique su respuesta.
6. Suponga que recibe una propuesta de investigación cuyo título es: "Nuevo modelo para abordar el estudio de las organizaciones empresariales". El autor del estudio dice que su trabajo se justifica porque es un nuevo esquema de analizar las organizaciones y, por tanto, es práctico. ¿Qué opina de este argumento para justificar el estudio?

www.FreeLibros.me

5.2. Justificación, objetivos e hipótesis de la investigación

5.2.1. Formulación de objetivos general y específicos

- Autor principal: **Fidias G. Arias**
- Libro recomendado: *El Proyecto de Investigación*
- Autor de apoyo: **César Bernal**
- Obra: *Metodología de la Investigación*

5.2.2. Elaboración de la justificación teórica, práctica y metodológica

- Autor principal: **César Augusto Bernal Torres**
- Libro recomendado: *Metodología de la Investigación*

5.2.3. Redacción de hipótesis de investigación e identificación de variables

- Autor principal: **Fred N. Kerlinger**
- Libro recomendado: *Investigación del Comportamiento*
- Autor de apoyo: **Roberto Hernández Sampieri**
- Obra: *Metodología de la Investigación*

5.2.1. Formulación de objetivos general y específicos

Para el desarrollo del subtema **Formulación de objetivos general y específicos**, se consideran aquellos contenidos relacionados con los **objetivos de investigación**, su formulación y la definición del objetivo general y los objetivos específicos, tomando como fuente principal a Fidias G. Arias y como autor de apoyo a César Augusto Bernal Torres. Se excluyen los contenidos relacionados con el planteamiento del problema, preguntas de investigación, justificación, hipótesis, metodología, técnicas de recolección, análisis de datos y conclusiones, debido a que corresponden a otros subtemas del programa.

N.°	Libros	Capítulo	Página	Contenido
1	*El Proyecto de Investigación* — **Fidias G. Arias**	**Capítulo 3. El problema de investigación**	**p. 43**	**3.4. Objetivos de investigación**
2	*Metodología de la investigación* — **César Augusto Bernal Torres**	**Capítulo 7. Proceso de investigación científica**	**p. 97**	**7.3. Objetivos de la investigación**

Arias ubica **"3.4. Objetivos de investigación"** en el capítulo dedicado al problema de investigación, mientras que Bernal presenta **"7.3. Objetivos de la investigación"** dentro del capítulo sobre el proceso de investigación científica.

Así quedaría la selección definitiva de fuentes para 5.2.1:
Arias → 3.4. Objetivos de investigación, p. 43
Bernal → 7.3. Objetivos de la investigación, p. 97.

Es importante señalar que algunos problemas no incluyen los tres elementos de una delimitación. El objeto de estudio determinará la pertinencia de incluirlos o no. Ejemplo:

¿Qué factores ocasionaron las pérdidas en la "Empresa HL" durante el período diciembre 2011 - febrero 2012?

En este caso se precisa el espacio: "Empresa HL", y el tiempo: período diciembre 2011 - febrero 2012. Sin embargo, no se hace referencia a una población específica, lo que también resulta válido.

3.4. Objetivos de investigación

Objetivo es sinónimo de meta, es decir, aquello que se aspira lograr o alcanzar. En este caso, nos referimos a meta en términos de conocimiento, es decir, los conocimientos que el investigador pretende obtener. En este orden de ideas:

> Objetivo de investigación es un enunciado que expresa lo que se desea indagar y conocer para responder a un problema planteado.

3.4.1. Características de los objetivos de investigación

a) Indican los conceptos que serán estudiados.
b) Precisan las variables o dimensiones que serán medidas.
c) Señalan los resultados que se esperan.
d) Definen los límites o alcances de la investigación.
d) Se redactan comenzando con un verbo en infinitivo.
e) Deben ser posibles de lograr.
f) Junto al problema de investigación, los objetivos responden a la pregunta ¿qué se pretende con la investigación? (Hernández y otros 2010; Ramírez, 2010; Sabino 2006), y no al ¿para qué?, como señalan algunos autores (ver ejemplo en la página 97).

A continuación se presenta una lista de verbos indicados para objetivos de investigación, clasificados según el nivel (Arias, 2006 b).

Nivel Exploratorio	Nivel Descriptivo	Nivel Explicativo
Conocer	Analizar	Comprobar
Definir	Calcular	Demostrar
Descubrir	Caracterizar	Determinar
Detectar	Clasificar	Establecer
Estudiar	Comparar	Evaluar
Explorar	Cuantificar	Explicar
Indagar	Describir	Inferir
Sondear	Diagnosticar	Relacionar
	Examinar	Verificar
	Identificar	
	Medir*	

Debe evitarse confundir los objetivos de investigación con:

a) Objetivos educativos o instruccionales: expresan lo que el alumno debe lograr como producto del proceso de enseñanza-aprendizaje.

Ejemplo:

"Al finalizar esta unidad, el alumno estará en capacidad de redactar oraciones completas..."

b) Objetivos prácticos: implican una habilidad psicomotora y son típicos de los proyectos tecnológicos que se desarrollan en áreas como Ingeniería, Electrónica y Computación.

Ejemplos:

"Diseñar un prototipo..."

"Desarrollar un programa computarizado (software)..."

*El verbo medir no debe ser entendido como la simple actividad que implica utilizar una cinta métrica para obtener la estatura de una persona, o las dimensiones de un terreno. En investigación social, medir significa la ejecución de objetivos de gran complejidad. Por ejemplo: medir la pobreza; medir la eficiencia, entre otros.

c) Metas empresariales: su logro depende de múltiples factores ligados a la organización, mas no de la voluntad del investigador.

Ejemplos:

"Aumentar la producción..."

"Incrementar las ventas..."

d) Propósitos: son ideas o intenciones cuya consecución puede escapar del alcance de la investigación.

Ejemplos:

"Motivar a los miembros de una comunidad..."

"Concientizar a los directivos..."

e) Actividades: son tareas o acciones implícitas en el proceso de investigación.

Ejemplos:

" Entrevistar a un grupo trabajadores..."

" Plantear una serie de recomendaciones..."

3.4.2. Tipos de objetivos de investigación

Los objetivos de investigación pueden ser generales o específicos.

Un objetivo general expresa el fin concreto de la investigación en correspondencia directa con la formulación del problema. Éste se puede descomponer, al menos, en dos objetivos específicos.

Mientras que:

Los objetivos específicos indican con precisión los conceptos, variables o dimensiones que serán objeto de estudio. Se derivan del objetivo general y contribuyen al logro de éste.

Ejemplo :

Objetivo general

Analizar las causas de la pobreza en Venezuela (1999-2005).

Objetivos específicos

1. Identificar la causas económicas de la pobreza en Venezuela.
2. Explicar las causas políticas que originan la pobreza en Venezuela.
3. Examinar las causas sociales: culturales y educativas, determinantes de la pobreza en Venezuela.

3.5. Correspondencia entre título, formulación del problema y objetivo general

La correspondencia entre título, formulación del problema y objetivo general radica en la presencia de elementos comunes, en los tres componentes del esquema de investigación.

Esta relación se presenta en los siguientes ejemplos:

Título	Formulación del problema	Objetivo general
Causas de la deserción escolar en la Educación Básica. Caso: Escuelas públicas del Distrito Metropolitano.	¿Cuáles son las causas de la deserción escolar en las Escuelas Básicas públicas del Distrito Metropolitano?	Identificar las causas de la deserción escolar en las Escuelas Básicas públicas del Distrito Metropolitano.
Impacto del control de cambio de divisas en el volumen de las importaciones venezolanas.	¿Qué impacto ocasionará el control de cambio de divisas en el volumen de las importaciones venezolanas?	Determinar el impacto del control de cambio de divisas en el volumen de las importaciones venezolanas.

7.3 Objetivos de la investigación

Un aspecto definitivo en todo proceso de investigación es la definición de los objetivos o del rumbo que debe tomar la investigación que va a realizarse. Así, los objetivos son los propósitos del estudio, expresan el fin que pretende alcanzarse; por tanto, todo el desarrollo del trabajo de investigación se orientará a lograr estos objetivos.

Éstos deben ser claros y precisos para evitar confusiones o desviaciones; sin embargo, esto no implica que los objetivos no puedan modificarse durante la realización de la investigación, porque en algunos casos hay que hacerlo.

7.3.1 Cómo se redactan y definen los objetivos en la investigación

Puesto que todo objetivo implica la acción que se desea lograr, es importante tener en cuenta que al redactar los objetivos de la investigación deben utilizarse verbos en infinitivo.

No es necesario escribir preámbulos al momento de redactar los objetivos; incluso se recomienda expresar directamente el objetivo. Por ejemplo, si un objetivo es "hacer un análisis de la situación actual del sector de las artes gráficas en la ciudad de...", no es necesario agregar frases previas al objetivo como; "debido a que las empresas del sector de las artes gráficas atraviesan una situación económica difícil, en este estudio se pretende hacer un análisis...".

Otro aspecto muy importante en el momento de plantear los objetivos de la investigación es utilizar verbos que puedan lograrse o alcanzarse durante el desarrollo de la investigación:

Los verbos podrían ser:

Determinar	Verificar	Definir
Identificar	Diseñar	Conocer
Evaluar	Elaborar	Estudiar
Describir	Proponer	Plantear
Formular	Analizar	Corroborar

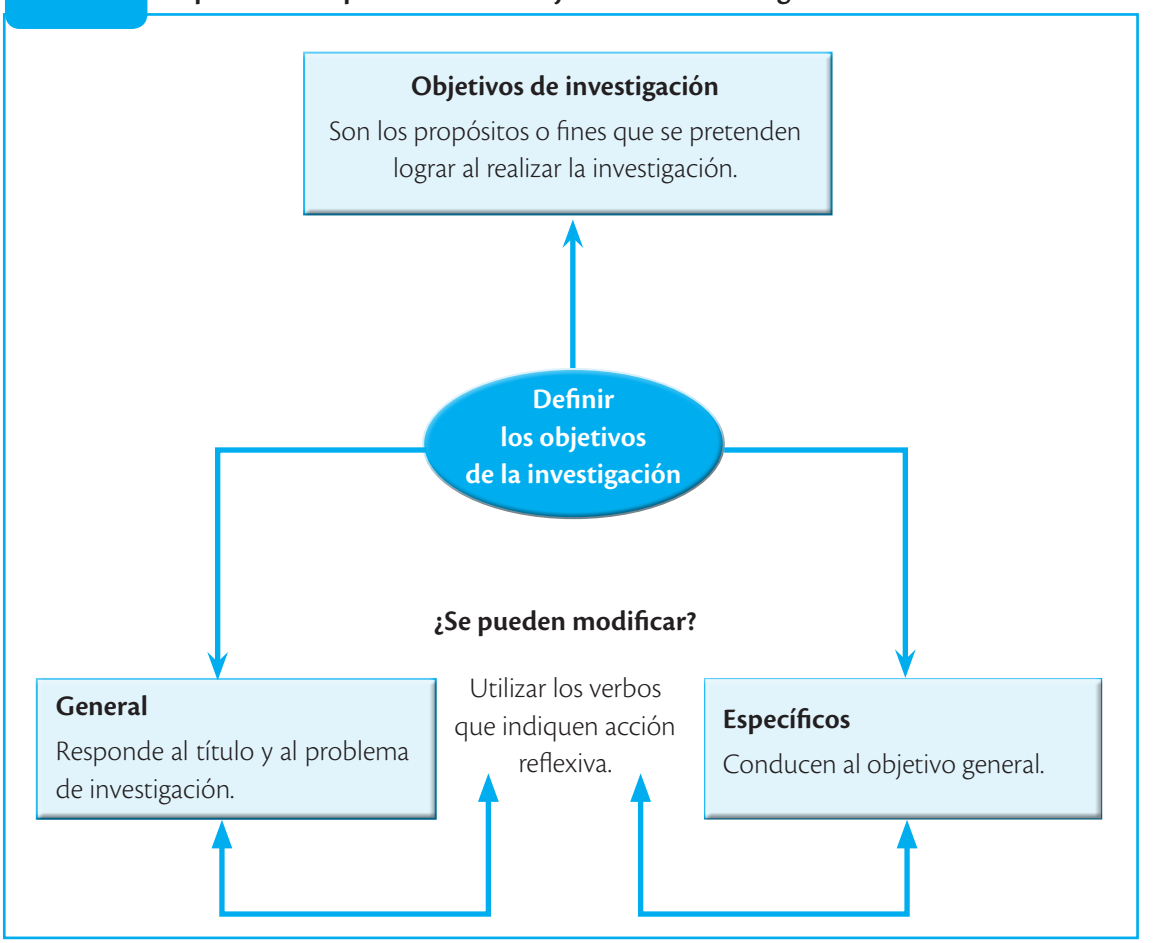
El uso de verbos como capacitar, cambiar, motivar, enseñar, mejorar y muchos otros que implican acciones finales debe ser prudente, porque estas acciones casi nunca se logran durante el progreso de la investigación, debido a que implican dedicarles tiempo y recursos y, muchas veces, tomar decisiones para desarrollar el objetivo propuesto.

www.FreeLibros.me

Sería deseable que todos los objetivos propuestos en un estudio pudieran llevarse a la práctica, con lo cual se contribuiría a solucionar problemas verdaderos. Sin embargo, el hecho de que no se apliquen no debe ser motivo de frustraciones, porque en el pregrado el solo hecho de realizar el ejercicio teórico de la investigación es ya un gran avance; y en el caso de la maestría, lo importante es la reflexión académica y fundamentada que pueda hacerse sobre el quehacer investigativo y la actitud que al respecto se cree.

La Figura 7.5 muestra los aspectos clave para la definición de los objetivos de la investigación, comenzando por un claro concepto de estos y la categorización en sus niveles de general y específicos, así como el tipo de verbos que han de utilizarse para su definición y la posibilidad de modificarse justificadamente.

FIGURA 7.5 **Aspectos clave para definir los objetivos de la investigación**



En los ejemplos 7.5 y 7.6 se muestran las limitaciones para alcanzar algunos objetivos.

www.FreeLibros.me

EJEMPLO 7.5

Si un estudiante en uno de sus objetivos se propone "Mejorar las condiciones ambientales de las empresas para generar un mejor entorno laboral", este objetivo implica realizar un diagnóstico de las condiciones ambientales actuales en las empresas; identificar dónde es necesario mejorar, proponer las estrategias para el mejoramiento de las condiciones ambientales y poner en marcha tales estrategias hasta producir el cambio o la mejora.

Esta última parte, que considera implementar las estrategias, casi nunca la aplica el estudiante; él sólo busca proponer estrategias para mejorar y hasta ahí llega su trabajo.

EJEMPLO 7.6

Al plantearse el objetivo "Disminuir el nivel de estrés en los miembros de familias y de sus miembros en enfermedad terminal", este es poco realizable durante el desarrollo de un trabajo de grado. Posiblemente el estudiante logre plantear estrategias para disminuir el grado de estrés, pero ello no significa que en realidad disminuya.

Para ello se requiere desarrollar un programa que implica tiempo y recursos, y el estudiante usualmente no llega a esa etapa en su trabajo de grado.

7.3.2 Objetivos: general y específicos

En toda investigación se requiere plantear dos niveles en los objetivos: El general y los específicos.

El objetivo general Debe reflejar la esencia del planteamiento del problema y la idea expresada en el título del proyecto de investigación.

Los objetivos específicos Se desprenden del general y deben formularse de forma que estén orientados al logro del objetivo general, es decir, que cada objetivo específico esté diseñado para lograr un aspecto de aquél; y todos en su conjunto, la totalidad del objetivo general. Los objetivos específicos son los pasos que se dan para lograr el objetivo general. Los ejemplos 7.7 y 7.8 muestran objetivos de investigación.

www.FreeLibros.me

EJEMPLO 7.7 **Objetivos en el campo de la economía y los negocios**

Título de la investigación: *Actitud de los directivos nacionales respecto a la innovación y la creatividad en sus empresas.*

Objetivos de la investigación:

Objetivo general

Analizar la actitud de los directivos nacionales de la industria manufacturera respecto a la innovación y creatividad en sus empresas.

Objetivos específicos

1. Elaborar un marco teórico sobre la actitud de los directivos de empresas respecto a la innovación y creatividad en las organizaciones productivas.
2. Diseñar un instrumento de medición de aptitudes con sus respectivas pruebas de validez y confiabilidad para aplicarlo a la población seleccionada.
3. Identificar la actitud de los gerentes de las 100 empresas más grandes del país respecto a la innovación y la creatividad en ellas.
4. Analizar los resultados obtenidos según subsector y demás variables sociodemográficas de las empresas estudiadas.

Título de la investigación: *Capacidad de adaptación de las empresas a los nuevos retos de competitividad y globalización.*

Objetivos de la investigación:

Objetivo general

Analizar la forma en que están preparándose las empresas nacionales para afrontar los retos de la competitividad y los demás cambios que caracterizan el nuevo ambiente de los negocios.

Objetivos específicos

1. Describir los retos que deben afrontar los negocios en el paradigma de la competitividad y la globalización.
2. Identificar las acciones que está realizando el gobierno para apoyar a las organizaciones ante los nuevos retos de la competitividad.
3. Conocer las acciones que han emprendido las organizaciones para hacerlas competitivas en un escenario combatiente y globalizado.
4. Evaluar qué papel desempeñan las instituciones universitarias en el proceso de preparación de organizaciones empresariales, para afrontar los retos de la nueva teoría económica y administrativa.

www.FreeLibros.me

Título de la investigación: *Impacto de las medidas económicas del gobierno en la actividad económica del país.*

Objetivos de la investigación:

Objetivo general

Evaluar el impacto de las medidas económicas del gobierno actual en la actividad empresarial.

Objetivos específicos

1. Describir las medidas económicas adoptadas por el gobierno en lo que va de su mandato.
2. Analizar el comportamiento de las tasas de interés, de la inflación y de las tasas de cambio durante el gobierno actual.
3. Identificar la opinión de los empresarios respecto a las medidas económicas del gobierno actual.
4. Identificar la opinión de los trabajadores sobre las medidas económicas del gobierno actual.
5. Analizar indicadores como el desempleo y la inversión social, y confrontarlas con la política económica del gobierno actual.

Título de la investigación: *Análisis comparativo entre el sistema de costos basado en actividades y el sistema de costeo tradicional.*

Objetivos de la investigación:

Objetivo general

Hacer un análisis comparativo entre el sistema de costeo basado en actividades y el sistema de costeo tradicional.

Objetivos específicos

1. Mostrar los fundamentos básicos del sistema de costeo basado en actividades y su diferencia con el sistema de costeo tradicional.
2. Evaluar las ventajas del sistema de costeo basado en actividades respecto al sistema de costeo tradicional.
3. Mostrar la importancia del costeo basado en actividades como herramienta de apoyo para la toma de decisiones gerenciales.

www.FreeLibros.me

EJEMPLO 7.8 **Objetivos de investigación en temas de psicología y educación**

Título de la investigación: *Análisis de la relación entre el tipo de valores practicados por los estudiantes universitarios y su rendimiento académico.*

Objetivos de la investigación:

Objetivo general

Analizar la relación entre el tipo de valores practicados por los estudiantes universitarios y su rendimiento académico.

Objetivos específicos

1. Identificar el tipo de valores practicados por los estudiantes universitarios.
2. Analizar el tipo de valores que practican los estudiantes universitarios de más alto rendimiento.
3. Conocer el tipo de valores que practican los estudiantes universitarios de bajo rendimiento.
4. Analizar la diferencia significativa entre los valores que practican los estudiantes universitarios de más alto rendimiento y los que practican los estudiantes de bajo rendimiento.
5. Evaluar la relación entre el tipo de valores practicados por los estudiantes objeto de estudio y el rendimiento académico de los mismos.

Título de la investigación: *Estudio sobre el estrés en la familia urbana en Calamba.*

Objetivos de la investigación:

Objetivo general

Analizar los factores generadores de estrés en la familia urbana calambana y las implicaciones que estos factores generan en el funcionamiento familiar.

Objetivos específicos

1. Identificar los factores (internos y del entorno) generadores de estrés en las familias calambanas.
2. Conocer las implicaciones del estrés en las familias estudiadas y en cada uno de sus miembros.
3. Analizar las estrategias que adoptan las familias calambanas para manejar el estrés.
4. Evaluar diferencias significativas entre los factores generadores de estrés y las estrategias utilizadas para el manejo de éste en las familias según estratos y nivel académico de sus miembros.

www.FreeLibros.me

Título de la investigación: *Estudio exploratorio sobre creatividad verbal, producción y comprensión de metáforas en niños de 4 a 12 años de edad.*

Objetivos de la investigación:

- ***Objetivo general***

Analizar la capacidad verbal, tanto para la producción como para la comprensión de metáforas, de niños y niñas de 4 a 12 años de edad, de estrato socioeconómico medio de la ciudad *XMJZO*.

- ***Objetivos específicos***

1. Evaluar la capacidad de producción de diferentes tipos de metáforas, en una muestra de niños y niñas de la ciudad *XMJZO*.
2. Evaluar la capacidad de comprensión de diferentes tipos de metáforas en la muestra de niños y niñas de la ciudad *XMJZO*.
3. Identificar la relación entre la capacidad de producción y de comprensión de diferentes tipos de metáforas en la muestra de niños y niñas que se van a evaluar.
4. Analizar diferencias y similitudes entre la capacidad de producción y comprensión de los diferentes tipos de metáforas, por parte de la muestra de niños y niñas de la ciudad *XMJZO*.
5. Discutir los resultados del trabajo de campo con los resultados de investigaciones realizadas por H. Gardner, citadas en el marco de referencia de este estudio.

Título de la investigación: *Evaluación de los efectos de un programa para la prevención de comportamientos agresivos de jóvenes en espectáculos masivos.*

Objetivos de la investigación:

- ***Objetivo general***

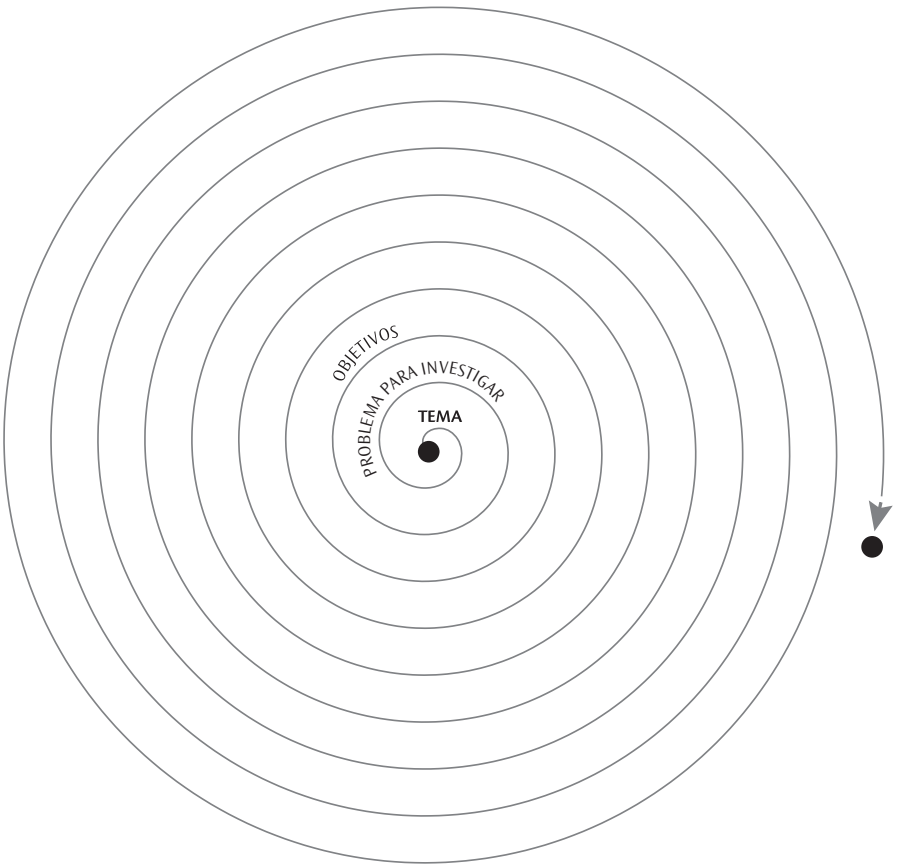
Evaluar los efectos de un programa de prevención de comportamientos agresivos de jóvenes que asisten a espectáculos masivos.

- ***Objetivos específicos***

1. Diseñar o adaptar un instrumento de evaluación, que sea válido para la medición del impacto del programa de prevención de comportamientos agresivos de parte de los jóvenes que asisten a espectáculos masivos.
2. Hacer el trabajo de evaluación de impacto.
3. Analizar la información obtenida en el trabajo de campo, orientada a evaluar el impacto del programa de prevención de los comportamientos agresivos de los jóvenes que asisten a espectáculos masivos.
4. Formular recomendaciones sobre los efectos del programa, así como evaluar y diseñar estrategias para la difusión del estudio, si los resultados de la evaluación son positivos.

www.FreeLibros.me

FIGURA 7.6 **El proceso de investigación. Establecimiento de los objetivos de la investigación**



RESUMEN

En el proceso de investigación, los objetivos definen el rumbo que debe tomar la investigación que va a realizarse; son los propósitos del estudio, y expresan el fin que pretende alcanzarse; el desarrollo del trabajo de investigación se orientará a lograr estos objetivos. No obstante, estos objetivos pueden modificarse durante la realización de la investigación, siempre que haya razones para hacerlo.

Los objetivos en investigación deben ser claros, precisos y realizables, y su redacción exige la utilización de verbos en infinitivo que implican acción reflexiva más que operativa. Estos objetivos se han de plantear en dos niveles: general y específico. El objetivo general debe reflejar la esencia del planteamiento del problema y la idea expresada en el título del proyecto de investigación. Los objetivos específicos se desprenden del general y deben formularse orientados al logro de éste, es decir, cada objetivo específico debe diseñarse para lograr un aspecto de aquél; y todos en su conjunto, la totalidad del objetivo general.

www.FreeLibros.me

5.2.2. Elaboración de la justificación teórica, práctica y metodológica

Para el desarrollo del subtema **Elaboración de la justificación teórica, práctica y metodológica**, se considera el apartado de **justificación de la investigación**, entendido como parte del proceso de elaboración de un proyecto de investigación. En la obra de César Augusto Bernal Torres, este contenido se encuentra desarrollado dentro del apartado **7.4. Justificación y delimitación de la investigación**, ubicado en la página 106. Para este subtema se abordará específicamente la **justificación**, sin desarrollar la delimitación de la investigación, debido a que no corresponde a la tarea asignada.

N.°	Libros	Capítulo	Página	Contenido
1	**Bernal Torres, César Augusto.** ***Metodología de la Investigación*****.**	**Capítulo 7: Proceso de investigación científica**	**p. 106**	**7.4. Justificación y delimitación de la investigación.** Para esta tarea se considera específicamente la parte correspondiente a la **justificación de la investigación**, excluyendo la delimitación.

7.4 Justificación y delimitación de la investigación

Toda investigación está orientada a la resolución de algún problema; por consiguiente, es necesario *justificar,* o exponer, los motivos que merecen la investigación. Asimismo, debe determinarse su cubrimiento o dimensión para conocer su viabilidad.

7.4.1 Criterios de justificación

De acuerdo con Méndez, la justificación de una investigación puede ser de carácter teórico, práctico o metodológico (Méndez, 1995).

Justificación teórica

En investigación hay una *justificación teórica* cuando el propósito del estudio es generar reflexión y debate académico sobre el conocimiento existente, confrontar una teoría, contrastar resultados o hacer epistemología del conocimiento existente.

En el caso de las ciencias económico-administrativas, un trabajo investigativo tiene justificación teórica cuando se cuestiona una teoría administrativa o una económica (es decir, los principios que la soportan), su proceso de implantación o sus resultados.

Cuando en una investigación se busca mostrar las soluciones de un modelo, está haciéndose una justificación teórica.

Si en una investigación se proponen nuevos paradigmas o se hace una reflexión epistemológica, se tiene una justificación eminentemente teórica, aunque al implementarla se vuelve práctica, ya que, como afirma López Cerezo (1988), toda investigación en alguna medida tiene la doble implicación, teórica y práctica.

La justificación teórica es la base de los programas de doctorado y algunos programas de maestría donde se tiene como objetivo la reflexión académica.

Justificación práctica

Se considera que una investigación tiene *justificación práctica* cuando su desarrollo ayuda a resolver un problema o, por lo menos, propone estrategias que al aplicarse contribuirían a resolverlo.

Los estudios de investigación de pregrado y de posgrado, en el campo de las ciencias económicas y administrativas, en general son de carácter práctico, o bien, describen o analizan un problema o plantean estrategias que podrían solucionar problemas reales si se llevaran a cabo.

Cuando en un trabajo de grado se realiza un análisis económico de un sector de la producción, su justificación es práctica porque genera información que podría utilizarse para tomar medidas tendientes a mejorar ese sector.

www.FreeLibros.me

Cuando un trabajo de grado se orienta a conocer los factores de motivación más utilizados en un determinado sector económico o en una empresa, su justificación es práctica, porque, al igual que en el caso del análisis del sector, la información sirve para actuar sobre la empresa, para mejorar o realizar *benchmarking* en otras organizaciones, o para confrontar la teoría sobre el tema.

Justificación metodológica

En investigación científica, la *justificación metodológica* del estudio se da cuando el proyecto que se va a realizar propone un nuevo método o una nueva estrategia para generar conocimiento válido y confiable.

7.4.2 Limitaciones del estudio o de la investigación

Una vez justificada la investigación, es necesario plantear las limitaciones dentro de las cuales se realizará (no todos los estudios tienen las mismas limitaciones, pues cada investigación es particular).

Las limitaciones en un proyecto de investigación pueden referirse a:

Limitaciones de tiempo Es necesario determinar cuál será el período, sea retrospectivo o prospectivo, dentro del cual se realizará el estudio del hecho, la situación, el fenómeno o población investigados. Por ejemplo, si el interés es estudiar el comportamiento de un sector económico, o los indicadores económicos de un país, es indispensable definir durante qué período (años, meses) se realizará tal análisis.

Limitaciones de espacio o territorio Son aquellas demarcaciones del espacio geográfico dentro del cual tendrá lugar una investigación. Las investigaciones pueden limitarse a una zona de una ciudad, a una ciudad, una región, un país, un continente, etcétera.

Limitaciones de recursos Mencionan la disponibilidad de recursos financieros para la realización del proyecto de investigación.

Además de estas limitaciones, en un proyecto de investigación puede haber limitaciones de información, población disponible para el estudio, dificultad de acceso a la misma, etcétera.

El investigador debe explicitar las limitaciones del proyecto con el propósito de facilitar su viabilidad.

EJEMPLO 7.9 **Enunciados de limitaciones de tiempo**

- Diagnóstico del sector de las flores en los últimos 12 años (1997-2009).
- Evaluación de la tasa de desempleo durante los últimos cuatro gobiernos.
- Análisis de las estrategias didácticas utilizadas en el proceso enseñanza-aprendizaje de la educación superior en los últimos 10 años.

www.FreeLibros.me

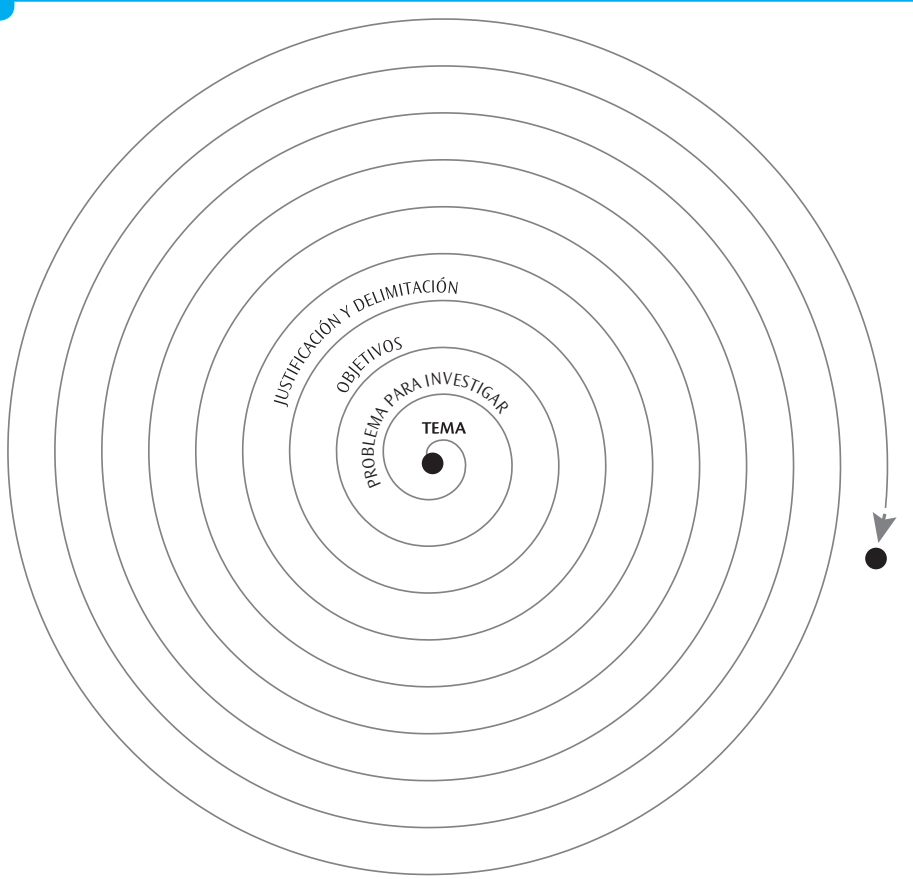
Existen investigaciones que se circunscriben a una determinada región, y eso se conoce como limitaciones geográficas.

EJEMPLO 7.10 **Enunciados de limitaciones de espacio o territorio**

- Cuando se dice que el estudio se realizará en empresas localizadas en la ciudad *PQTL* o en las empresas latinoamericanas, se hace referencia a una limitación geográfica.
- Afirmar que un estudio se realizará en una determinada ciudad o zona geográfica implica delimitar este estudio geográfica o territorialmente.

En general, puede hablarse de varias limitaciones según las características de la población o muestra sobre la cual recae el estudio. Estas características se detallarán cuando se haga referencia a la población y muestra objeto de estudio.

FIGURA 7.7 **El proceso de investigación. Justificación y delimitación de la investigación**



www.FreeLibros.me

La Figura 7.7 sintetiza la justificación como las razones de carácter práctico, teórico y metodológico que se tienen para realizar el estudio, así como sus delimitaciones en los ámbitos espacial, cronológico y/o sociodemográfico.

RESUMEN

En investigación, la justificación se refiere a las razones del porqué y el para qué de la investigación que se va a realizar, es decir, justificar una investigación consiste en exponer los motivos por los cuales es importante llevar a cabo el respectivo estudio. Al respecto, suele haber tres dimensiones o tipos de justificación: teórica, práctica y metodológica. Algunas investigaciones pueden requerir los tres tipos de justificación, otras dos tipos y otras solo uno. Ello depende de las particularidades de cada investigación.

La delimitación o el alcance en investigación se refiere a la dimensión o al cubrimiento que ésta tendrá en el espacio geográfico, período de tiempo y perfil sociodemográfico del objeto de estudio.

EJERCICIOS DE REPASO Y ANÁLISIS

1. ¿Qué significa justificar y delimitar una investigación?
2. Explique las distintas formas de justificar una investigación.
3. Seleccione un tema de investigación y justifique por qué debe realizarse el estudio.
4. Revise los ejemplos de los temas propuestos en el tema relacionado con la formulación de objetivos que justifiquen la realización de esos estudios.
5. Suponga que quiere realizar un estudio sobre el concepto de ser humano que han manejado las distintas teorías administrativas a través del tiempo. ¿Cuál es la justificación más adecuada para este estudio? Justifique su respuesta.
6. Suponga que recibe una propuesta de investigación cuyo título es: "Nuevo modelo para abordar el estudio de las organizaciones empresariales". El autor del estudio dice que su trabajo se justifica porque es un nuevo esquema de analizar las organizaciones y, por tanto, es práctico. ¿Qué opina de este argumento para justificar el estudio?

www.FreeLibros.me

5.2.3. Redacción de hipótesis de investigación e identificación de variables

Para el desarrollo del subtema **Redacción de hipótesis de investigación e identificación de variables**, se consideran aquellos contenidos relacionados directamente con la **formulación, características e importancia de las hipótesis**, así como con el **concepto, identificación y clasificación de las variables** dentro de la investigación científica. Se priorizan los contenidos de los capítulos 2 y 3 de *Investigación del Comportamiento*, evitando aquellos relacionados con el planteamiento del problema, la metodología, la recolección de datos, la medición y el análisis estadístico, debido a que corresponden a otros subtemas del esquema.

N.°	Libros	Capítulo	Página	Contenido
1	*Investigación del Comportamiento* — Fred N. Kerlinger y Howard B. Lee	**Capítulo 2. Problemas e hipótesis**	**p. 23**	**Hipótesis:** concepto y función de las hipótesis dentro de la investigación científica.
2	*Investigación del Comportamiento* — Fred N. Kerlinger y Howard B. Lee	**Capítulo 2. Problemas e hipótesis**	**p. 24**	**Importancia de los problemas e hipótesis:** importancia de la hipótesis para orientar la investigación.
3	*Investigación del Comportamiento* — Fred N. Kerlinger y Howard B. Lee	**Capítulo 2. Problemas e hipótesis**	**p. 25**	**Virtudes de los problemas e hipótesis:** características que contribuyen a la utilidad científica de una hipótesis.
4	*Investigación del Comportamiento* — Fred N. Kerlinger y Howard B. Lee	**Capítulo 2. Problemas e hipótesis**	**p. 28**	**Generalidad y especificidad de los problemas e hipótesis:** relación entre el grado de generalidad y especificidad de una hipótesis.
5	*Investigación del Comportamiento* — Fred N. Kerlinger y Howard B. Lee	**Capítulo 2. Problemas e hipótesis**	**p. 29**	**Naturaleza multivariable de la investigación y problemas del comportamiento:** consideración de múltiples variables en la formulación de hipótesis.
6	*Investigación del Comportamiento* — Fred N. Kerlinger y Howard B. Lee	**Capítulo 3. Constructos, variables y definiciones**	**p. 36**	**Variables:** concepto de variable y su función dentro de la investigación.

7	*Investigación del Comportamiento* — Fred N. Kerlinger y Howard B. Lee	**Capítulo 3. Constructos, variables y definiciones**	**p. 42**	**Tipos de variables:** clasificación de las variables utilizadas en la investigación.
8	*Investigación del Comportamiento* — Fred N. Kerlinger y Howard B. Lee	**Capítulo 3. Constructos, variables y definiciones**	**p. 42**	**Variables independientes y dependientes:** identificación y diferenciación de las variables según su función en la investigación.
9	*Investigación del Comportamiento* — Fred N. Kerlinger y Howard B. Lee	**Capítulo 3. Constructos, variables y definiciones**	**p. 47**	**Variables continuas y categóricas:** diferenciación de estos tipos de variables.
10	*Investigación del Comportamiento* — Fred N. Kerlinger y Howard B. Lee	**Capítulo 3. Constructos, variables y definiciones**	**p. 48**	**Constructos observables y variables latentes:** distinción entre estos conceptos en la investigación.

de la forma de las relaciones antes planteada. Sin embargo, en la mayoría de los problemas de investigación del comportamiento, se estudian las relaciones entre dos o más variables; por ello, enfatizaremos ese tipo de enunciados de relación.

Criterios de los problemas y enunciados de problemas

Existen tres criterios de buenos problemas y enunciados de problemas. El primero: el problema debe expresar una relación entre dos o más variables. En efecto plantea preguntas como la siguiente: ¿*A* está relacionada con *B*? ¿Cómo están relacionadas *A* y *B* con *C*? ¿Cómo está *A* relacionada con *B* bajo las condiciones *C* y *D*? La excepción a esta consideración ocurre casi siempre en investigación metodológica o taxonómica.

Segundo: el problema debe ser enunciado de manera clara y sin ambigüedades en forma de pregunta. En lugar de decir, por ejemplo "el problema es..." o "El propósito de este estudio es..." resulta necesario plantear una pregunta. Las preguntas tienen la virtud de presentar los problemas directamente. El propósito de un estudio no es por fuerza el mismo que el problema de un estudio. El propósito del estudio de Hurlock, por ejemplo, fue arrojar luz sobre el uso de incentivos en las situaciones escolares. El problema consistió en la pregunta acerca de la relación entre incentivos y rendimiento. Otra vez, mientras más simple, mejor: formule una pregunta.

El tercer criterio con frecuencia es difícil de satisfacer. Demanda que el problema y su enunciado *impliquen* la posibilidad de ser sometidos a una prueba empírica. Un problema que no contenga implicaciones para probar las relaciones que enuncia, no constituye un problema científico. Esto significa no sólo que se enuncie una relación real, sino también que las variables de la relación puedan ser medidas de alguna forma. Hay muchas preguntas interesantes e importantes que no constituyen preguntas científicas tan sólo porque no son susceptibles de prueba. Ciertas preguntas filosóficas y teológicas, aunque importantes para quienes las consideran, no pueden ser probadas empíricamente, por lo que no generan interés para el científico como tal. La pregunta epistemológica "¿Cómo conocemos?" es una pregunta de ese tipo. La educación plantea muchas preguntas interesantes pero no científicas, por ejemplo "¿Mejora la educación democrática el aprendizaje de los jóvenes?" "¿Son buenos los procesos grupales para los niños?" Estas preguntas pueden ser etiquetadas como metafísicas en el sentido en que están, al menos así enunciadas, fuera de la posibilidad de una prueba empírica. Las principales dificultades estriban en que algunas no constituyen relaciones, y es muy difícil o imposible definir la mayoría de sus constructos de forma que puedan ser medidos.

Hipótesis

Una *hipótesis* es un enunciado conjetural de la relación entre dos o más variables. Las hipótesis siempre se presentan en forma de enunciados declarativos y relacionan, de manera general o específica, variables con variables. Hay dos criterios que definen a las "buenas" hipótesis y a sus enunciados. Son los mismos que mencionamos para los problemas y sus enunciados. 1) Las hipótesis son enunciados acerca de las relaciones entre variables. 2) Las hipótesis contienen implicaciones claras para probar las relaciones enunciadas. Estos criterios significan que los enunciados de hipótesis contienen dos o más variables, que son medibles o pueden serlo, y que especifican cómo están relacionadas las variables.

Permítanos mencionar tres hipótesis de la literatura y aplicarles estos criterios. La primera hipótesis procede de un estudio de Wegner y colaboradores. (1987) que parece desafiar el sentido común: a mayor supresión de pensamientos indeseados, mayor preocu-

pación por ellos (represión ahora; obsesión más tarde). Aquí se establece una relación entre una variable, supresión de una idea o pensamiento, y otra variable, preocupación u obsesión. Dado que ambas se definen y miden con facilidad, las implicaciones para probar la hipótesis también se conciben sin esfuerzo. Los criterios están satisfechos. En el estudio de Wegner y colaboradores, se les pidió a los sujetos que *no* pensaran en un "oso blanco". Cada vez que pensaran en él, debían de tocar una campana. El número de campanadas indicaba el nivel de preocupación. Una segunda hipótesis, que resulta inusual y corresponde al estudio de Ayres y Hughes (1986), enuncia la relación de una forma que llamamos nula: el nivel de ruido o música no tiene efecto en el funcionamiento visual. La relación se establece con claridad: una variable, intensidad del sonido (por ejemplo, música), se relaciona con otra, funcionamiento visual, a través de las palabras "no tiene efecto en". En el criterio de potencia de ser probada, sin embargo, encontramos dificultades. Nos enfrentamos con el problema de definir "funcionamiento visual" e "intensidad" de forma que puedan medirse. Si podemos resolver este problema de manera satisfactoria, entonces, tenemos en definitiva una hipótesis. Ayres y Hughes lo hicieron al definir intensidad como 107 decibeles y funcionamiento visual en términos de una puntuación en una tarea de agudeza visual. Esta hipótesis permitió contestar una pregunta que la gente con mucha frecuencia se hace: "¿por qué bajamos el volumen del estéreo del auto cuando *buscamos* una dirección?". Ayres y Hughes encontraron una caída marcada en el funcionamiento perceptual cuando el nivel de música llegaba a 107 decibeles.

La tercera hipótesis representa una categoría numerosa e importante. En ella la relación es indirecta, oculta. En general enuncia que los grupos *A* y *B* diferirán en alguna característica. Por ejemplo: las mujeres creen, con mayor frecuencia que los hombres, que deben perder peso aun cuando éste se encuentre dentro de los límites normales (Fallon y Rozin, 1985). Esto es, que las mujeres difieren de los hombres en cuanto a la percepción de su figura corporal. Observe que este enunciado está un paso más allá de la hipótesis real que puede plantearse como: la percepción de la figura corporal es, en parte, una función del género. Si los enunciados posteriores constituyeran la hipótesis enunciada, entonces la primera podría llamarse una subhipótesis o una predicción específica basada en la hipótesis original.

Consideremos otra hipótesis de este tipo pero dando un paso más adelante. Los individuos que tienen características iguales o similares tendrán actitudes similares hacia objetos cognitivos significativamente relacionados con su papel ocupacional (Saal y Moore, 1993). (*Los objetos cognitivos* se definen como algo concreto o abstracto, percibido y "conocido" por los individuos. Personas, grupos, ascenso en el trabajo o en las calificaciones, el gobierno y la educación son algunos ejemplos.) La relación en este caso es, por supuesto, entre características personales y actitudes (hacia un objeto cognitivo relacionado con la característica personal, por ejemplo, género y actitudes hacia otros que reciben una promoción). Para probar esta hipótesis, sería necesario tener al menos dos grupos, cada uno con una característica diferente, y después comparar las actitudes de ambos grupos. Por ejemplo, como en el caso del estudio de Saal y Moore, la comparación sería entre hombres y mujeres. Serían comparados en relación a su evaluación hacia el ascenso dado a un compañero de trabajo del mismo sexo o del opuesto. En este ejemplo, se satisfacen los criterios.

Importancia de los problemas e hipótesis

Hay poca duda de que las hipótesis son herramientas importantes e indispensables de la investigación científica. Existen tres razones principales para esta creencia. La primera es que son, digamos, los instrumentos de trabajo de la teoría. Las hipótesis pueden deducirse

a partir de la teoría y de otras hipótesis. Si por ejemplo, trabajamos en una teoría sobre la agresividad, se presume que buscamos causas y efectos del comportamiento agresivo. Es posible que hayamos observado casos de comportamiento agresivo ocurrido después de circunstancias frustrantes. La teoría, entonces, puede incluir la proposición: la frustración produce agresividad (Berkowitz, 1983; Dill y Anderson, 1995; Dollard, Doob, Miller, Mowrer, y Sears, 1939). A partir de esta amplia hipótesis, podemos deducir hipótesis más específicas, como por ejemplo: impedir que los niños alcancen sus metas deseadas (frustración) generará pleitos entre ellos (agresión); si los niños son privados del amor paterno (frustración), reaccionarán en parte, con un comportamiento agresivo.

La segunda razón es que es posible someter a prueba las hipótesis y demostrar que son probablemente verdaderas o probablemente falsas. No se prueban hechos aislados, como se dijo antes, sólo relaciones. Es probable que la principal razón de usar hipótesis en la investigación científica sea que constituyen proposiciones relacionales. En esencia, son predicciones del tipo: "si *A*, entonces *B*", que utilizamos para probar la relación entre *A* y *B*. Dejamos que los hechos tengan la oportunidad de establecer la probable veracidad o falsedad de la hipótesis.

La tercera razón es que las hipótesis son herramientas poderosas para el avance del conocimiento porque permiten al científico ir más allá de sí mismo. Aunque desarrolladas por humanos, las hipótesis existen, pueden ser probadas y puede demostrarse que son probablemente correctas o incorrectas de manera independiente a los valores y opiniones de una persona (sesgos). Esto resulta crítico: no habría ciencia, en sentido completo alguno, sin las hipótesis.

Tan importantes como las hipótesis son los problemas tras ellas. Como Dewey (1938/1982, pp. 105-107) ha señalado, la investigación por lo general empieza con un problema. Indica que primero hay una situación indeterminada en la que las ideas son vagas, aparecen dudas, y el pensador queda perplejo. Agrega que el problema no se enuncia; de hecho, no puede ser enunciado hasta que uno ha experimentado una situación tan indeterminante.

La indeterminación, sin embargo, deberá, en última instancia ser eliminada. Aunque es cierto, como se señaló antes, que el investigador con frecuencia puede tener sólo una noción general y difusa del problema, tarde o temprano habrá de definir una idea clara de lo que el problema es. Aunque este enunciado parezca obvio, una de las cosas más difíciles de lograr, es enunciar de una manera clara y completa el problema de investigación. En otras palabras, uno debe saber qué es lo que trata encontrar. Cuando por fin se identifica, el problema ya está en camino a la solución.

Virtudes de los problemas e hipótesis

Los problemas y las hipótesis tienen virtudes importantes: 1) dirigen la investigación (las relaciones expresadas en las hipótesis indican al investigador lo que debe hacer); 2) los problemas e hipótesis, dado que son de ordinario enunciados relacionales generalizados, permiten al investigador deducir manifestaciones empíricas específicas implicadas en ellos. Podemos decir, de acuerdo con Guida y Ludlow (1989): si es un hecho verdadero que los niños de un tipo de cultura (Chile) tienen un mayor grado de ansiedad que los niños de otro tipo de cultura (blancos estadounidenses), entonces los niños en la cultura chilena deben rendir menos en lo académico que los niños en la cultura estadounidense. Los niños chilenos quizá también debieran presentar una menor autoestima o un *locus* de control más externo en lo que se refiere a la escuela y a la labor académica.

Hay diferencias importantes entre problemas e hipótesis. Las hipótesis, si están enunciadas de manera apropiada, pueden ser probadas. Una hipótesis dada puede ser demasia-

do amplia para ser probada de forma directa; sin embargo, si es una "buena" hipótesis, es posible deducir a partir de ella otras que sí lo sean. Los hechos o las variables no se prueban como tales. Se prueban las relaciones enunciadas por las hipótesis. Un problema no puede ser resuelto de manera científica a menos que se reduzca a su forma de hipótesis, ya que un problema es una pregunta, generalmente de naturaleza amplia, que no puede probarse en forma directa. No se someten a prueba preguntas como: ¿la presencia o ausencia de otra persona en un sanitario público afecta la higiene personal? (Pedersen, Keithly y Brady, 1986). ¿Las sesiones de consejería grupal disminuyen el nivel de morbilidad psiquiátrica en oficiales de policía? (Doctor, Cutris e Issacs, 1994). Quizás uno probaría una o más hipótesis deducidas de estas preguntas. Por ejemplo, para estudiar el último problema, uno puede hipotetizar que los oficiales de policía que asisten a sesiones de consejería para reducir el estrés requerirán menos días de incapacidad por enfermedad que aquéllos que no asisten. La hipótesis para el primer problema podría indicar que la presencia de una persona en un sanitario público hará que otras se laven las manos.

Los problemas e hipótesis permiten que avance el conocimiento científico al ayudar al investigador a confirmar o refutar una teoría. Suponga que un investigador en el campo de la psicología aplica a unos sujetos tres o cuatro pruebas, entre las cuales hay una para evaluar la ansiedad relacionada con una prueba aritmética. Al calcular de manera rutinaria las correlaciones entre las tres o cuatro pruebas, uno encuentra que la correlación entre ansiedad y aritmética es negativa. De lo anterior se deduce que a mayor ansiedad, menor puntuación en la prueba de aritmética. Sin embargo, es muy probable que esta relación sea fortuita e incluso espuria, pero si el investigador hubiera hipotetizado la relación con base en una teoría, tendría mayor confianza en sus resultados. El investigador que no hipotetiza relaciones en forma previa, no permite que los hechos prueben o rechacen nada. Las palabras *probar* y *rechazar* no deben tomarse en su sentido literal: una hipótesis nunca se prueba o refuta realmente. Para ser más precisos, deberíamos decir algo del tipo de: el peso de la evidencia está del lado de la hipótesis o el peso de la evidencia arroja dudas sobre la hipótesis. Braithwaite (1953/1996, p. 14) dice:

> De este modo, la evidencia empírica nunca prueba la hipótesis: en casos apropiados podemos decir que *se establece* (itálicas agregadas) la hipótesis, lo que significa que la evidencia hace que sea razonable aceptar la hipótesis; pero ésta nunca *prueba* la hipótesis en el sentido de que la hipótesis sea una consecuencia lógica de la evidencia.

Este uso de la hipótesis es similar a participar en un juego de azar. Se establecen las reglas del juego y se definen las apuestas por adelantado. Uno no puede cambiar las reglas después de un resultado, como tampoco se pueden cambiar las apuestas una vez hechas. Uno no tira los dados primero y luego apuesta. No sería "justo". De igual forma, si en primera instancia se recolectan datos y después se toma uno de ellos y se llega a una conclusión con base en él, se han violado las reglas del juego científico. El juego no es "justo" porque el investigador puede capitalizar fácilmente, digamos, dos relaciones importantes de las cinco a prueba. Las otras tres, por lo general, se olvidan. En un juego "justo", se cuenta cada tiro del dado, en el sentido de que se gana o no con base en el resultado de cada tirada.

Las hipótesis dirigen la investigación. Como Darwin señaló hace más de 100 años, todas las observaciones han de ser a favor o en contra de algún punto de vista para tener alguna utilidad. Las hipótesis incorporan aspectos de la teoría bajo prueba de forma susceptible o casi susceptible de ser probada. Antes se dio un ejemplo de la teoría del reforzamiento en el que se dedujeron hipótesis demostrables a partir del problema general. Podemos demostrar la importancia del reconocimiento de esta función de las hipótesis al introducirnos por la puerta trasera y usar una teoría muy difícil o quizás imposible de probar. La teoría de Freud de la ansiedad incluye el constructo de la represión. Con este

término Freud se refería a la introducción forzada de ideas inaceptables en lo profundo del inconsciente. Para probar la teoría freudiana de la ansiedad es necesario deducir relaciones sugeridas por la teoría. Estas deducciones por fuerza deberán incluir el concepto de represión que implica el constructo del inconsciente. Es posible formular hipótesis que utilizan estos constructos; para probar la teoría han de ser formuladas de esta manera. Pero probarlos resulta más difícil debido a la extrema dificultad para definir términos como "represión" e "inconsciente" de manera que puedan medirse. Hasta hoy nadie ha tenido éxito al definir estos dos constructos sin apartarse en gran medida del significado y uso freudianos originales. Las hipótesis constituyen, entonces, puentes importantes entre la teoría y la investigación empírica.

Problemas, valores y definiciones

Para establecer con claridad la naturaleza de los problemas y de las hipótesis, analizaremos ahora dos o tres errores comunes. En primer término, los problemas científicos no son preguntas morales ni éticas: ¿Son las medidas disciplinarias de tipo punitivo perjudiciales para los niños? ¿Debiera ser el liderazgo de una organización de tipo democrático? ¿Cuál es la mejor forma de enseñar a los estudiantes universitarios? Formular estas preguntas equivale a presentar cuestionamientos de valor y juicio que la ciencia no puede contestar. Muchas de las que se han llamado hipótesis no lo son en absoluto. Por ejemplo: el método de enseñanza a pequeños grupos es mejor que el método expositivo. Éste es un enunciado de valor; constituye un acto de fe, no una hipótesis. Si fuera posible establecer una relación entre las variables, y definirlas de manera que se pudiera probar esa relación, entonces podríamos contar con una hipótesis. Pero no hay una forma científica de someter a prueba preguntas de valor.

Una forma rápida y relativamente fácil de detectar preguntas y enunciados de valor consiste en buscar palabras como *debe, debería, mejor que* (en lugar de *mayor que*), así como palabras similares que indiquen juicios culturales o personales, o preferencias (sesgos). Sin embargo, los enunciados de valor son engañosos. Aunque resulta obvio que un enunciado que incluye la palabra "debería" es un enunciado de valor, otros tipos no son tan evidentes. Consideremos el enunciado: los métodos autoritarios de enseñanza conducen a un pobre aprendizaje. En este caso sí hay una relación. Pero el enunciado falla como una hipótesis científica en tanto que incorpora dos expresiones de valor: "métodos autoritarios de enseñanza" y "pobre aprendizaje", ninguna de las cuales puede definirse con propósitos de medición sin borrar las palabras *autoritario* y *pobre*.[2]

Con frecuencia se formula otra clase de enunciados que no constituyen hipótesis o que son hipótesis pobres, en especial en el campo de la educación. Consideremos, por ejemplo, el siguiente: los cursos de tronco común representan una experiencia enriquecedora. Otro tipo de enunciado que se usa con frecuencia, es la generalización vaga: es posible identificar las habilidades de lectura en el segundo grado; La meta del individuo auténtico es la autorrealización; El prejuicio se relaciona con ciertos rasgos de personalidad.

Otro defecto común de los enunciados de problema aparece a menudo en las tesis doctorales: enlistar aspectos metodológicos o "problemas" como subproblemas. Estos aspectos metodológicos poseen dos características que hacen fácil detectarlos: 1) No son

[2] Un caso ya casi clásico del uso de la palabra *autoritario* es la frase que a veces se escucha entre educadores: el método expositivo es autoritario. Esto parece indicar que quien lo dice no gusta del método expositivo y lo considera negativo. De forma similar, una de las formas más efectivas de criticar a un maestro es decir que es autoritario.

problemas sustantivos que surgen del problema básico; y 2) Se relacionan con técnicas o métodos de muestreo, medición o análisis. En general, no aparecen en forma de pregunta y contienen palabras tales como *probar, determinar, medir*: "para *determinar* la confiabilidad de los instrumentos usados en esta investigación;" "para *probar* la significación de las diferencias entre las medias" o "para *asignar* alumnos de manera aleatorizada a los grupos experimentales", son ejemplos de esta noción equivocada de problemas y subproblemas.

Generalidad y especificidad de los problemas e hipótesis

Una dificultad que el investigador por lo general encuentra y que casi todos los estudiantes que trabajan en una tesis hallan molesta, es la generalidad y la especificidad de los problemas e hipótesis. Si el problema es muy general, es demasiado vago para ser sometido a prueba. Así, desde el punto de vista científico resulta inútil, aunque puede ser interesante para leer. Los problemas e hipótesis demasiado generales o vagos son comunes. Por ejemplo: La creatividad es una función de la autorrealización del individuo; La educación democrática potencia el aprendizaje social y la forma cívica; El autoritarismo en el salón de clases universitario inhibe la imaginación creativa del estudiante. Todos resultan problemas interesantes, pero en su forma actual son por completo inútiles en el terreno científico, en tanto que no pueden ser sometidos a prueba y porque parecen sugerir una seguridad espuria de que constituyen hipótesis que "algún día" pueden ser probadas.

Términos tales como "creatividad", "autorrealización", "democracia" y "autoritarismo" no tienen, al menos hasta ahora, referentes empíricos adecuados.[3] Es cierto que podemos definir *creatividad*, en una forma limitada al especificar una o dos pruebas de creatividad. Éste puede ser un procedimiento legítimo; sin embargo, al emplearlo, corremos el riesgo de alejarnos del término original y de su significado. Esto es en particular cierto cuando hablamos de creatividad artística. Desde luego, con frecuencia aceptamos el riesgo con tal de investigar problemas importantes. Aun así, términos como "democracia" son casi imposibles de definir. Incluso cuando lo hacemos, a menudo descubrimos que hemos destruido su significado original. Una excepción sobresaliente es la definición y la medición de "democracia" de Bollen (1980). Examinaremos ambas en otros capítulos.

El otro extremo es caer en demasiada especificidad. Todo estudiante ha escuchado que es necesario reducir los problemas a una dimensión manejable. Esto es cierto, pero por desgracia, podemos reducirlo tanto hasta hacerlo desaparecer. En general, mientras más específicos son el problema o la hipótesis, más claras resultan sus implicaciones a probar. Sin embargo, el precio que podemos pagar es la trivialidad. Los investigadores no pueden manejar problemas demasiado amplios por su tendencia a ser demasiado vagos en cuanto a las operaciones adecuadas de investigación. Por otro lado, en su entusiasmo por reducir el problema a un tamaño manejable o por encontrar un problema manipulable, pueden terminar con su vida, y convertirlo en trivial o carente de importancia. Por ejemplo, una tesis sobre la simple relación entre velocidad de lectura y tamaño de la letra, por muy interesante e importante que pudiera parecer, resulta débil para un estudio doctoral. El estudiante de ese nivel necesitará ampliar el tema al recomendar una comparación entre géneros y considerar variables como cultura y antecedentes familiares. El investigador podría también expandir el estudio para concentrarse en los niveles de iluminación y

[3] Aunque se ha conducido con éxito una variedad de estudios sobre autoritarismo, no es claro que entendamos lo que significa autoritarismo en el salón de clases. Por ejemplo, una acción de un maestro autoritario en un salón de clases puede no serlo en otra aula. La mencionada conducta democrática exhibida por un maestro puede ser etiquetada como autoritarismo si la muestra otro docente. Tal elasticidad no pertenece a la ciencia.

el tipo de letra. Demasiada especificidad quizás sea más dañina que demasiada generalidad. El investigador puede estar en posición de contestar una pregunta específica pero no podrá generalizar los hallazgos a otras situaciones o grupos de personas. A cualquier precio, alguna clase de compromiso debe establecerse entre generalidad y especificidad. La capacidad para definir tal compromiso de manera efectiva es, en parte, función de la experiencia, y en parte, del estudio crítico de los problemas de investigación.

He aquí algunos ejemplos de problemas de investigación contrastantes, ya sea muy generales o muy específicos:

1)	Demasiado general:	Existen diferencias de género al jugar.
	Demasiado específico:	La puntuación de Carlos será 10 puntos mayor que la de Carol en el juego profesional Tetris.
	Cerca del ideal:	Habrá mayor transferencia de aprendizaje al practicar videojuegos con niños que con niñas.
2)	Demasiado general:	Las personas pueden leer letras de mayor tamaño más rápido que las letras más pequeñas.
	Demasiado específica:	Los alumnos de último año de la escuela Duarte pueden leer tipos de letra de 24 puntos más rápido que tipos de letra de 12 puntos.
	Cerca del ideal:	Una comparación de tres diferentes tamaños de letra y agudeza visual en la velocidad de lectura y de comprensión.

La naturaleza multivariable de la investigación y problemas del comportamiento

Hasta este punto, la discusión de problemas e hipótesis se ha limitado a dos variables, x y y. Debemos corregir cualquier impresión de que tales problemas e hipótesis son la norma en la investigación del comportamiento. Los investigadores en psicología, sociología, educación y otras ciencias del comportamiento se han concientizado de la naturaleza multivariable de este tipo de estudios. En lugar de decir: si p, entonces q, es frecuente y más apropiado decir: si $p_1, p_2 \ldots, p_k$, entonces q; o bien: si p, entonces q, bajo las condiciones r, s y t.

A continuación, un ejemplo que puede aclarar este punto. En lugar de simplemente formular la hipótesis: si hay frustración, entonces hay agresividad, es más realista reconocer la naturaleza multivariable de los determinantes e influencias de la agresividad. Esto se logra al decir, por ejemplo: si se es muy inteligente, de clase media, varón y frustrado, entonces hay agresividad; o bien: si hay frustración, entonces hay agresividad bajo las condiciones de alta inteligencia, clase media y sexo masculino. En lugar de tener una x, nosotros ahora tenemos cuatro x. Aunque un fenómeno puede ser el más importante para determinar o influir en otro fenómeno, es poco probable que la mayoría de los fenómenos de interés para los científicos del comportamiento sean determinados de forma simple. Es mucho más probable que lo sean de manera múltiple. Es mucho más probable que la agresividad sea el resultado de diversas influencias que actúan de forma compleja. Más aún, la agresividad en sí misma contiene múltiples aspectos. Después de todo, hay diferentes clases de agresividad.

Los problemas y las hipótesis, entonces deben reflejar la complejidad multivariable de la realidad psicológica, sociológica y educativa. Hablaremos de una x y una y, en particular en la parte inicial de este libro. Sin embargo, *es preciso* entender que la investigación del

comportamiento, que tuvo un enfoque casi exclusivamente univariado, se ha tornado cada vez más multivariable. Nos hemos propuesto usar la palabra "multivariable" en lugar de "multivariada" por una razón importante. De forma tradicional los estudios "multivariados" son aquéllos que tienen más de una variable *y* y una o más variables *x*. Cuando hablamos de una variable *y* y más de una variable *x* se usa el término "multivariable" que resulta más apropiado para hacer la distinción. Por ahora usaremos "univariado" para indicar una *x* y una *y*. De manera estricta, el término "univariado" también se aplica a *y*. Pronto encontraremos conceptos y problemas de naturaleza multivariada. Secciones posteriores del libro estarán enfocadas en especial a un enfoque y énfasis de este tipo. Para más explicaciones sobre las diferencias entre *multivariable* y *multivariado* (véase Kleinbaum, Kupper, Muller y Nizam, 1997).

Comentarios finales: el poder especial de las hipótesis

A veces se oye decir que las hipótesis son innecesarias en la investigación. Algunos sienten que las hipótesis restringen innecesariamente su imaginación investigadora y que el trabajo de la ciencia y de la investigación científica es descubrir cosas nuevas y no elaborar lo obvio. Algunos piensan que las hipótesis son obsoletas. Tales afirmaciones resultan engañosas y malinterpretan el propósito de las hipótesis.

Casi puede decirse que las hipótesis son uno de los instrumentos más poderosos que ha inventado el hombre para alcanzar un conocimiento confiable. Observamos un fenómeno; especulamos sobre sus causas posibles. Naturalmente, nuestra cultura tiene respuestas para dar cuenta de la mayoría de los fenómenos —muchas correctas, muchas incorrectas, muchas una mezcla de hechos y supersticiones, y muchas pura superstición—. Es obligación del científico dudar de la mayor parte de las explicaciones sobre los fenómenos. Esas dudas son sistémicas. Los científicos insisten en someter las explicaciones sobre los fenómenos a una prueba empírica controlada. Para lograrlo, formulan las explicaciones en términos de teorías e hipótesis. De hecho, las explicaciones constituyen hipótesis. Los científicos sólo disciplinan la cuestión al escribirla en forma de hipótesis sistemáticas y comprobables. Si una explicación no puede formularse en términos de una hipótesis comprobable, deberá considerarse como una explicación metafísica y, por lo tanto, no susceptible de investigación científica. Como tal, los científicos la rechazan como carente de interés.

El poder de las hipótesis va más allá. La hipótesis constituye una predicción: indica que si ocurre *x*, también ocurrirá *y*; esto es, *y* se predice a partir de *x*. Entonces si se hace que *x* ocurra (es decir, que varíe) y se observa que *y* también ocurre (o sea, varía de forma concomitante), entonces la hipótesis se confirma. Resulta una evidencia más poderosa que la simple observación, sin predicción, la covariación de *x* y *y*. Es más poderosa en el sentido de apuesta-juego antes discutido. El científico apuesta a que *x* conduce hacia *y*. Si en un experimento, *x* conduce en efecto a *y*, entonces habrá ganado la apuesta. Una persona no puede sólo entrar en el juego en cualquier momento y observar una ocurrencia común quizá fortuita de *x* y *y*. No se juega de esta forma (al menos no en nuestra cultura). La persona debe jugar de acuerdo con las reglas, y las reglas en la ciencia están hechas para minimizar el error y la falibilidad. Las hipótesis forman parte de las reglas del juego científico.

Aun cuando no se confirmen las hipótesis, tienen poder. Aun cuando *y* no covaríe con *x*, el conocimiento avanza. Los hallazgos negativos en ocasiones resultan tan importantes como los positivos, puesto que reducen el universo total de la ignorancia, y algunas veces señalan hacia otras hipótesis y líneas de investigación. *Pero el científico no puede distinguir la evidencia positiva de la negativa hasta usar una hipótesis.* Por supuesto, es posible conducir

Conceptos y constructos

Los términos "concepto" y "constructo" tienen significados similares, aunque existe una diferencia importante. Un "concepto" expresa una abstracción creada por una generalización a partir de instancias particulares. "Peso" es un concepto que expresa numerosas observaciones de cosas que son "más o menos" y "pesadas o ligeras". "Masa", "energía" y "fuerza" son conceptos usados por científicos de la física. Por supuesto, son mucho más abstractos que conceptos como "peso", "alto" y "longitud".

Un concepto de mayor interés para los lectores es el "aprovechamiento". Es una abstracción que se genera a partir de la observación de ciertos comportamientos de los niños, que se asocian con el dominio del "aprendizaje" en tareas escolares —lectura de palabras, resolución de problemas aritméticos, elaboración de dibujos, etcétera—. Los diversos comportamientos observados se reúnen y expresan en una palabra. "Aprovechamiento", "inteligencia", "agresividad", "conformidad" y "honestidad" son conceptos usados para expresar la variedad del comportamiento humano.

Un *constructo* es un concepto, que tiene el significado agregado de haber sido inventado o adoptado para un propósito científico especial, de forma deliberada y consciente. "Inteligencia" es un concepto, una abstracción de la observación de comportamientos presumiblemente inteligentes y no inteligentes. Sin embargo, como todo constructo científico, "inteligencia" implica tanto más como menos de lo que pueda significar como concepto. Esto quiere decir que los científicos de manera consciente y sistemática la usan en las dos formas: 1) se incorpora en los esquemas teóricos y se relaciona en diversas formas con otros constructos (podemos decir, por ejemplo, que el aprovechamiento escolar es en parte función de la inteligencia y motivación) y 2) "inteligencia" se define y especifica de tal forma que pueda ser observada y medida (podemos hacer observaciones de la inteligencia de los niños al aplicar una prueba de inteligencia, o al solicitar a los maestros que señalen el grado relativo de inteligencia de sus alumnos).

Variables

Los científicos de forma algo vaga, llaman a los constructos o propiedades que estudian, "variables". Algunos ejemplos de variables importantes en sociología, psicología, ciencia política y educación son: género, ingreso, educación, clase social, productividad organizacional, movilidad ocupacional, nivel de aspiración, aptitud verbal, ansiedad, afiliación religiosa, preferencias políticas, desarrollo político (de las naciones), orientación ocupacional, prejuicios raciales y étnicos, conformidad, recuerdo, memoria de reconocimiento y aprovechamiento. Puede decirse que una variable es una propiedad que asume diversos valores. Siendo redundantes, una variable es algo que varía. Aunque esta forma de expresarlo nos aporta una noción intuitiva de lo que son, necesitamos una definición al mismo tiempo más general y precisa.

Una *variable* es un símbolo al que se le asignan valores o números. Por ejemplo, x es una variable: es un *símbolo* al que se le asignan valores numéricos. La variable x puede tomar cualquier conjunto justificable de valores, por ejemplo, puntajes en una prueba de inteligencia o en una escala de actitudes. En el caso de la inteligencia, asignamos a x un conjunto de valores numéricos proporcionados por el procedimiento especificado en una determinada prueba de inteligencia. Este grupo de valores varía de bajo a alto, por ejemplo de 50 a 150.

Una variable, x, sin embargo, puede tener sólo dos valores. Si el género es el constructo bajo estudio, entonces a x se le pueden asignar 1 y 0, donde 1 representa uno de los géne-

ros y 0 el otro. Aún así, es una variable. Otros ejemplos de variables con dos valores son: dentro–fuera, correcto–incorrecto, viejo-joven, ciudadano-no ciudadano, clase media-clase trabajadora, maestro-no maestro, republicano-demócrata, etcétera. Tales variables se llaman *dicotomías*, variables dicotómicas o binarias.

Algunas de las variables usadas en la investigación conductual son verdaderas dicotomías, es decir, se caracterizan por la presencia o ausencia de una propiedad: masculino-femenino, con hogar-indigente, empleado-desempleado. Otras variables son *politomías*. Un buen ejemplo es la preferencia religiosa: protestante, católico, musulmán, judío, budista, otra. Tales dicotomías y politomías se denominan "variables cualitativas". La naturaleza de este calificativo se analizará más adelante. En teoría, muchas variables, sin embargo, pueden asumir valores continuos. Ha sido una práctica común en la investigación del comportamiento convertir las variables continuas en dicotómicas o politómicas. Por ejemplo, la inteligencia, una variable continua, se ha dividido en inteligencia alta, media y bajas, o en alta y baja. Variables como ansiedad, introversión y autoritarismo han recibido un trato similar. Aunque no es posible convertir una variable que de manera natural es dicotómica, como género, en una variable continua, siempre podemos convertir una variable continua en una dicotómica o politómica. Más adelante se verá que tal conversión puede tener un propósito conceptual útil, pero para el análisis de datos constituye una práctica negativa en tanto que descarta información.

Definiciones constitutivas y operacionales de constructos y variables

La diferencia que hicimos antes entre "concepto" y "constructo" conduce de manera natural a otra distinción importante entre los tipos de definiciones de constructos y variables. Se puede definir a las palabras o constructos de dos formas generales: primero, podemos definir una palabra con el uso de otras palabras, que es lo que hace un diccionario. Podemos definir *inteligencia* como "un intelecto operante", "agudeza mental" o "la habilidad para pensar de forma abstracta". Tales definiciones utilizan otros conceptos o expresiones conceptuales en lugar de la expresión o palabra que se define. Segundo, podemos definir una palabra por las acciones o comportamientos que expresa o implica. Definir *inteligencia* de esta forma requiere que especifiquemos qué comportamientos de los niños son "inteligentes" y cuáles son "no inteligentes". Podemos decir que un niño de siete años de edad que lee una historia con éxito es "inteligente", si el niño no puede leer la historia podemos asumir que el chico "no es inteligente". En otras palabras, esta clase de definición puede llamarse *definición observacional* o *conductual*. Se usan definiciones a partir de "otras palabras" y definiciones "observacionales" de manera cotidiana.

En esta discusión hay una imprecisión perturbadora. Aunque los científicos utilizan los tipos de definiciones que acabamos de describir, lo hacen en una forma más precisa. Expresamos este uso al definir y explicar la diferencia que Margenau (1950/1977) plantea para las definiciones constitutivas y operacionales. Una definición *constitutiva* define un constructo usando otros constructos. Por ejemplo, podemos definir *peso* diciendo que es la "pesadez" de los objetos, o *ansiedad* como "temor subjetivo". En ambos casos hemos sustituido un concepto por otro. Algunos de los constructos de una teoría científica pueden definirse de manera constitutiva. Torgerson (1958/1985), a partir de las ideas de Margenau, indica que para ser útiles desde el punto de vista científico, todos los constructos deben poseer significado constitutivo, es decir, poder ser usados en teorías.

Una definición *operacional* asigna significado a un constructo o variable al especificar las actividades u "operaciones" necesarias para medirlo y evaluar la medición. De manera

alternativa, una definición operacional constituye una especificación de las actividades del investigador para medir una variable o para manipularla. Implica algo así como un manual de instrucciones para el investigador. En efecto, dice, "haga tal y cual, de la forma tal y tal". En síntesis, define o aporta significado a una variable al delinear paso a paso lo que el investigador debe hacer para medirla y para evaluar dicha medición.

Michel (1990) presenta una excelente revisión histórica de cómo las definiciones operacionales se hicieron populares en las ciencias sociales y del comportamiento. Michel cita a P. W. Bridgeman, premio Nobel, como creador de la definición operacional en 1927. Bridgeman, como lo relata Michel (1990, p. 15), indica: "en general, con cualquier concepto dado queremos significar tan sólo un conjunto de operaciones; *el concepto es sinónimo del correspondiente conjunto de operaciones*". Cada operación diferente definiría un concepto distinto.

Un ejemplo bien conocido, aunque extremo, de una definición operacional es: inteligencia (ansiedad, aprovechamiento, etcétera) es la puntuación en una prueba *X* de inteligencia, o inteligencia es lo que la prueba *X* de inteligencia mide. Las puntuaciones altas indican un mayor nivel de inteligencia que las bajas. Esta definición indica qué hacer para medir la inteligencia y no precisa qué tan bien es medida la inteligencia por el instrumento especificado. (Se presume que se indagó sobre la adecuación de esta prueba antes de que el investigador la usara.) En este tipo de uso, una definición operacional equivale a una ecuación en la que decimos: "si la inteligencia es igual a la puntuación en la prueba *X* de inteligencia, las puntuaciones altas indican un mayor grado de inteligencia que los bajos". Parecería también que estamos diciendo "el significado de inteligencia (en este estudio) se expresa por el puntaje en la prueba *X* de inteligencia".

Existen, en general, dos clases de definiciones operacionales: 1) las *medidas* y 2) las *experimentales*. La definición de arriba está más estrechamente ligada con las definiciones medidas que con las experimentales. Una definición operacional *medida* describe cómo será medida una variable. Por ejemplo, aprovechamiento puede definirse por una prueba estandarizada de aprovechamiento, por un examen desarrollado por el maestro, o por las calificaciones. Doctor, Cutris e Isaacs (1994), al estudiar el efecto de la consejería para el estrés en oficiales de policía, definieron de manera operacional morbilidad psiquiátrica como las puntuaciones en el Cuestionario General de Salud y el número de días que habían tomado por incapacidad. Altas puntuaciones y un gran número de días indicaban niveles elevados de morbilidad. Little, Sterling y Tingstrom (1996) estudiaron los efectos de la raza y origen geográfico en la atribución. La atribución se definió operacionalmente como la puntuación en el cuestionario de estilo atribucional. Un estudio puede incluir la variable *consideración* que puede definirse operacionalmente a través de una lista de comportamientos de niños que se presume son comportamientos considerados. Después, se puede pedir a los maestros que evalúen a los alumnos en una escala de cinco puntos. Tales comportamientos pueden ejemplificarse como instancias en que un niño le dice a otro "lo siento" o "discúlpame", o cuando un chico le presta un juguete a otro o cuando se lo pide (sin la amenaza de agresión), o cuando un pequeño ayuda a otro con una tarea escolar. También puede definirse por la suma de comportamientos considerados: a mayor cantidad, mayor nivel de consideración.

Una definición operacional *experimental* señala los detalles (operaciones) de la manipulación de una variable por parte del investigador. El reforzamiento puede definirse operacionalmente al precisar los detalles de cómo los sujetos serán reforzados (premiados) y no reforzados (no premiados) por comportamientos específicos. Hom, Berger, Duncan, Miller y Belvin (1994) definieron operacionalmente el reforzamiento en forma experimental. En este estudio, los niños fueron asignados a uno de cuatro grupos. Dos de los grupos estuvieron sujetos a una condición de reforzamiento con base en cooperación,

mientras que los otros dos trabajaron en un esquema en que se reforzaba la postura individualista. Bahrick (1984) definió la memoria a largo plazo en términos de al menos dos procesos en lo referente a la retención de información de tipo académico. Un proceso llamado "almacenaje permanente" (*permastore*), elige de manera selectiva alguna información para ser almacenada de forma permanente y resulta muy resistente al decaimiento (olvido). El otro proceso al parecer elige información menos significativa por lo que es menos resistente al olvido. Esta definición contiene implicaciones claras para la manipulación experimental. Strack, Martin y Stepper (1988) definieron operacionalmente la sonrisa como la activación de los músculos asociados con la sonrisa humana. Lo lograron al solicitar que una persona sostuviera una pluma en su boca de cierta manera. Se trata de un procedimiento no intrusivo en tanto que no se les pidió a los participantes que posaran con una sonrisa. Se presentarán otros ejemplos de ambos tipos de definiciones más adelante.

Los investigadores científicos eventualmente enfrentan la necesidad de medir las variables de las relaciones que estudian. Algunas mediciones son fáciles, otras difíciles. Medir el género o la clase social es fácil; pero evaluar creatividad, conservadurismo o efectividad organizacional resulta difícil. La importancia de las definiciones operacionales no puede dejar de enfatizarse. Ellas son ingredientes indispensables de la investigación científica porque permiten al investigador medir variables y porque representan puentes entre el nivel de la teoría-hipótesis-constructo y el nivel de observación. No hay investigación científica sin observaciones, y éstas no son posibles sin instrucciones claras y específicas de qué y cómo observar. Las definiciones operacionales son tales instrucciones.

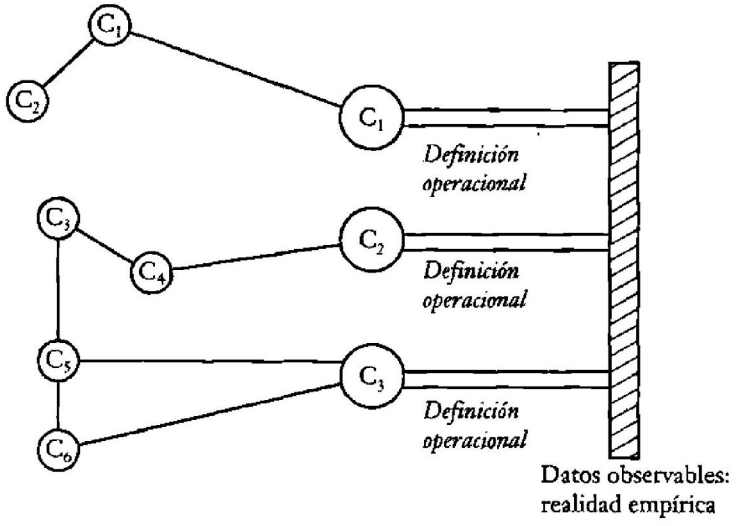
Aunque indispensables, las definiciones operacionales sólo aportan significados limitados de los constructos. Ninguna definición operacional puede expresar toda la riqueza y los diversos aspectos de algunas variables, como sucede con el prejuicio humano. Esto implica que las variables medidas por los científicos siempre tienen un significado limitado y específico. La "creatividad" que estudian los psicólogos no se refiere necesariamente a la "creatividad" de los artistas, aunque, por supuesto, tengan elementos comunes. Una persona que piensa en una solución creativa para un problema matemático puede mostrar una escasa creatividad como poeta (Barron y Harrington, 1981). Algunos psicólogos han definido operacionalmente a la creatividad como el desempeño en la prueba de Torrance de Pensamiento Creativo (Torrance, 1982). Los niños que obtienen una puntuación alta en esta prueba, tienen mayor probabilidad de exhibir logros creativos en la edad adulta.

Algunos científicos afirman que estos limitados significados operacionales son los únicos significados que "significan" algo, que todas las demás definiciones son disparates metafísicos. Señalan que las discusiones sobre ansiedad constituyen tonterías de tipo metafísico, a menos que se cuente con y se usen, definiciones operacionales adecuadas. Éste es un punto de vista extremo, aunque posee algunos aspectos saludables. Insistir en que cada término que usemos en el discurso científico sea definido operacionalmente sería demasiado reduccionista y restrictivo y, como se verá, científicamente erróneo. Northrop (1947/1983, p. 130) señala, por ejemplo: "La importancia de las definiciones operacionales estriba en que hacen posible la verificación y enriquecen el significado. Sin embargo, no agotan el significado científico". Margenau (1950/1977, p. 232) señala el mismo punto en su extensa discusión sobre los constructos científicos.

A pesar de los riesgos del operacionalismo extremo, parece seguro decir que ha sido y es una influencia saludable. Como señala Skinner (1945, p. 274):

> La actitud operacional, a pesar de sus inconvenientes, es positiva en cualquier ciencia, pero en especial en psicología, por la presencia de un vasto vocabulario de origen antiguo y no científico.

FIGURA 3.1



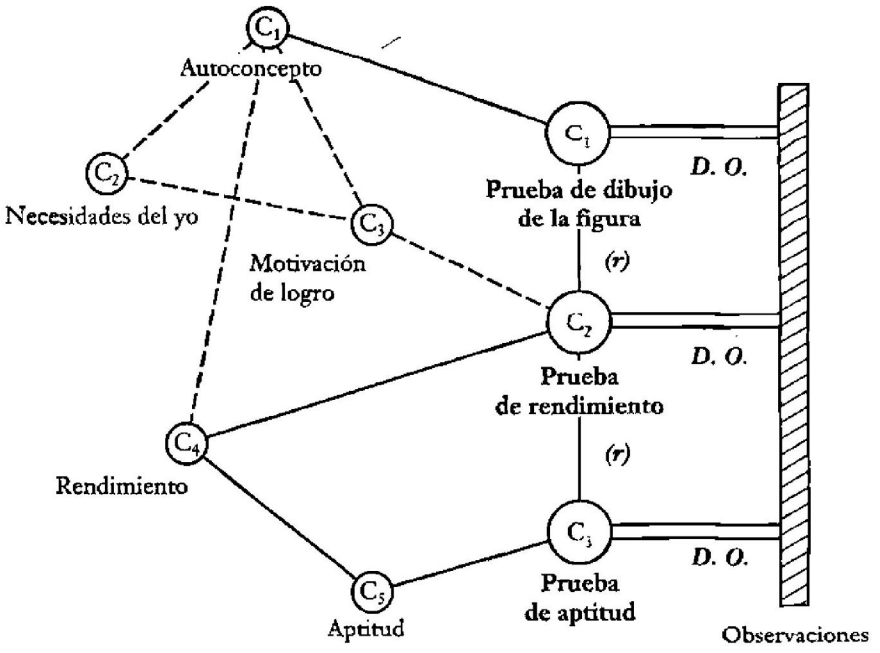
Cuando se consideran los términos usados en educación, es claro que esta disciplina también posee un vasto vocabulario de origen antiguo y no científico. Consideremos éstos: el niño integral, el enriquecimiento horizontal y vertical, la satisfacción de las necesidades del alumno, el tronco común, el ajuste emocional y el enriquecimiento curricular. Esto también se aplica al campo de la atención geriátrica. Aquí la enfermera especializada maneja términos como el proceso de envejecimiento, la autoimagen, el mantenimiento de la atención y la negligencia unilateral (Eliopoulos, 1993; Smeltzer y Bare, 1992).

Para aclarar las definiciones constitutivas y operacionales (así como la teoría) veamos la figura 3.1, que ha sido adaptada de Margenau (1950/1977) y de Torgerson (1958/1985). El diagrama intenta ilustrar una teoría bien desarrollada. Las líneas sencillas representan conexiones teóricas o relaciones entre constructos. Estos constructos, etiquetados con letras minúsculas, se definen de manera constitutiva; esto es, c_4 está definida de alguna forma por c_3 o viceversa. Las líneas dobles representan definiciones operacionales. Los constructos están ligados de forma directa a datos observables, y son vinculaciones indispensables con la realidad empírica. Sin embargo, no todos los constructos en una teoría científica se definen operacionalmente. De hecho, la teoría que tiene todos sus constructos así definidos es algo tenue.

Construyamos una "pequeña teoría" del bajo rendimiento para ilustrar estos conceptos. Supongamos que un investigador cree que el bajo rendimiento es en parte una función del autoconcepto de los alumnos. Piensa que los estudiantes que se perciben como inadecuados y que tienen percepciones negativas de sí mismos, también tienden a rendir menos que lo que su capacidad potencial y aptitudes indican. De aquí se sigue que las necesidades del yo (que no definiremos aquí) y la motivación de logro (que llamaremos necesidad de logro) están ligadas al bajo rendimiento. Como es natural, el investigador está consciente de la relación entre aptitud e inteligencia y logro en general. Un diagrama que ilustra esta "teoría" puede ser como el de la figura 3.2.

El investigador no cuenta con una medida *directa* de autoconcepto, pero supone que puede inferirlo a partir de una prueba de dibujo de la figura. Entonces define operacional-

FIGURA 3.2



mente el autoconcepto como ciertas respuesta a esa prueba. Éste es probablemente el método más común para medir constructos psicológicos (y educativos). La línea gruesa entre c_1 y C_1 indica la naturaleza relativamente directa de la relación que se asume entre el autoconcepto y la prueba. (La línea doble entre C_1 y el nivel de observación indica una definición operacional, como en la figura 3.1.)

De forma parecida, el constructo de rendimiento (c_4) se define operacionalmente como la discrepancia entre la medición realizada del rendimiento (C_2) y la de la aptitud (c_5). En este modelo el investigador no mide directamente la motivación de logro, ni tampoco posee una definición operacional de ella. En otro estudio se puede hipotetizar específicamente una relación entre rendimiento y motivación de logro, en cuyo caso se tratará de definir motivación para el logro en forma operacional.

Una línea continua sencilla entre conceptos, por ejemplo la que está entre el constructo rendimiento (c_4) y prueba de aprovechamiento (C_2), indica una relación relativamente bien establecida entre el rendimiento postulado y lo que miden las pruebas estandarizadas de rendimiento. La líneas continuas sencillas entre C_1 y C_2 y aquellas entre C_2 y C_3 indican relaciones obtenidas entre las puntuaciones de las pruebas de estas mediciones. (Las líneas entre C_1 y C_2, y entre C_2 y C_3 están marcadas como (r) para significar "relación" o "coeficiente de correlación".)

Las líneas discontinuas representan relaciones postuladas entre constructos que no están relativamente bien establecidas. Un buen ejemplo de esto es la relación postulada entre el autoconcepto y la motivación de logro. Uno de los propósitos de la ciencia es convertir estas líneas punteadas en continuas al cerrar la brecha entre definición-opera-

cional-medición. En este caso, es concebible que tanto el autoconcepto como la motivación de logro puedan ser definidas operacionalmente y medidas directamente.

En esencia, ésta es la forma en que el científico del comportamiento opera. Este especialista se desplaza de manera continua entre el nivel del constructo y el de la observación y lo logra al definir operacionalmente las variables de la teoría que pueden serlo. Luego, se estiman las relaciones entre la definición operacional y las variables medidas. De estas relaciones estimadas, el científico abstrae inferencias acerca de las relaciones entre los constructos. En el ejemplo anterior, el científico del comportamiento calcula la relación entre C_1 (la prueba de dibujo de la figura) y C_2 (prueba de rendimiento). Si la relación se establece en este nivel observacional, el científico infiere que existe una relación entre c_1 (autoconcepto) y c_4 (rendimiento).

Tipos de variables

Variables independientes y dependientes

Dejemos atrás los fundamentos de las definiciones para regresar a las variables. Es posible clasificar las variables de diversas formas. En este libro, tres tipos de variables son muy importantes y serán enfatizadas: 1) variables independientes y dependientes, 2) variables activas y atributo, y 3) variables continuas y categóricas.

La forma más útil de categorizar las variables es como independientes o dependientes. Esta taxonomía resulta muy útil debido a su aplicabilidad general, su simplicidad y su importancia especial tanto en la conceptualización como en el diseño de la investigación, así como en la comunicación de los resultados de ésta. Una *variable independiente* es la causa *supuesta* de la *variable dependiente*, el efecto *supuesto*. La variable independiente es el antecedente; la dependiente es el consecuente. Dado que uno de los objetivos de la ciencia es descubrir relaciones entre diferentes fenómenos, la búsqueda de las relaciones entre variables independientes y dependientes lo logra. Se asume que la variable independiente influye en la dependiente. En algunos estudios, la variable independiente "causa" cambios en la variable dependiente. Cuando decimos: si A, entonces B, tenemos una conjunción condicional de una variable independiente (A) y una variable dependiente (B).

Los términos "variable independiente" y "variable dependiente" proceden de las matemáticas, donde X es la variable independiente y Y es la dependiente. Ésta es probablemente la mejor forma de pensar en las variables independientes y dependientes porque no hay necesidad de utilizar la discutida palabra *causa* y las palabras afines a ella, y dado que el uso de tales símbolos se aplica a la mayoría de las situaciones de investigación. No hay restricción teórica alguna en cuanto a la cantidad de X y Y. Cuando más adelante consideremos el pensamiento y análisis multivariado, trataremos con diversas variables dependientes e independientes.

En los experimentos, el investigador manipula la variable independiente. Los cambios en los valores o niveles de la variable independiente generan cambios en la variable dependiente. Cuando investigadores del campo educativo estudiaron los efectos de diferentes métodos de enseñanza en el desempeño en una prueba de matemáticas, ellos variaron los métodos de enseñanza. En una condición pudieron tener "sólo exposición frontal", en la otra pudo haber "exposición frontal y video". El método de enseñanza es la variable independiente. La variable resultado, la puntuación en la prueba de matemáticas, es la variable dependiente.

La asignación de participantes a diferentes grupos con base en la existencia de alguna característica es un ejemplo de cuando el investigador no puede manipular la variable

independiente. En esta situación, los valores de la variable independiente son preexistentes. El participante tiene la característica o no la tiene. En este caso, no hay posibilidad de una manipulación experimental, pero se considera que "lógicamente" la variable tiene algún efecto en la variable dependiente. Las variables de características del sujeto constituyen la mayor parte de este tipo de variables independientes. Una de las variables independientes más comunes de este tipo es el género (femenino y masculino). Así, si un investigador desea determinar si hombres y mujeres difieren en las destrezas matemáticas, se aplicaría una prueba matemática a representantes de ambos grupos, y se compararían las puntuaciones de la prueba. La prueba de matemáticas sería la variable dependiente. Una regla general es que cuando el investigador manipula una variable o asigna participantes a los grupos según alguna característica, esa variable es la independiente. La tabla 3.1 muestra una comparación entre los dos tipos de variables independientes y su relación con la variable dependiente. La variable independiente debe tener al menos dos niveles o valores. Observe en la tabla 3.1 que ambas situaciones presentan dos niveles para la variable independiente.

La variable dependiente es, por supuesto, *hacia* la que se hace la predicción, mientras que la independiente es aquella *a partir de* la cual se predice. La variable dependiente, Y, es el efecto supuesto, que varía de manera concomitante a los cambios o variaciones en la variable independiente, X; es la variable que se observa para detectar variaciones como un resultado supuesto de la variación en la variable independiente. La variable dependiente es el resultado medido que el investigador usa para determinar si los cambios en la variable independiente tuvieron un efecto. Al predecir Y a partir de X, podemos tomar cualquier valor de X que deseemos, mientras que el valor de Y que predecimos *es* "dependiente" del valor de X que hemos elegido. En general, la variable dependiente es la condición que tratamos de explicar. Por ejemplo, la variable dependiente más común en educación, es "aprovechamiento" o "aprendizaje". Deseamos explicar o dar cuenta del aprovechamiento. Para ello tenemos un gran número de posibles X o variables independientes de dónde elegir.

Cuando se estudia la relación entre inteligencia y aprovechamiento escolar, la inteligencia es la variable independiente y el aprovechamiento es la dependiente. (¿Se podría concebir a la inversa?) Otras variables independientes que pueden estudiarse con relación a la variable dependiente aprovechamiento son: clase social, métodos de enseñanza, tipos de personalidad, tipos de motivación (recompensa y castigo), actitudes hacia la escuela y ambiente en el salón de clases, entre otros. Cuando se estudian los supuestos determinantes de la delincuencia, aquellos tales como condiciones de pobreza, hogares desintegrados, falta de amor de los padres y aspectos similares, constituyen las variables independientes y,

TABLA 3.1 *Relación de variables independientes manipuladas y no manipuladas con la variable dependiente*

Niveles de la variable independiente

Método de enseñanza		**Género**	
Sólo exposición frontal Puntuaciones en la prueba de matemáticas	Exposición frontal y video Puntuaciones en la prueba de matemáticas	Femenino Puntuaciones en la prueba de matemáticas	Masculino Puntuaciones en la prueba de matemáticas
Variables dependientes		Variables dependientes	

como es natural, la delincuencia (o mejor aún, el comportamiento delictivo) es la variable dependiente. En la hipótesis frustración-agresión, frustración es la variable independiente, y agresión, la dependiente. En ocasiones, un fenómeno se estudia por sí mismo y ya sea la variable dependiente o la independiente están implícitas. Es el caso en que se estudian los comportamientos y características del maestro. La variable dependiente implícita común es el aprovechamiento o el comportamiento del niño, aunque el comportamiento del maestro puede, por supuesto, constituir una variable dependiente. Consideremos un ejemplo en el campo de la atención médica. Cuando se comparan medidas cognitivas y funcionales de pacientes con Alzheimer entre instituciones de internamiento tradicional y unidades de cuidado especial, la variable independiente es el lugar de atención. Las variables dependientes son las medidas cognitivas y funcionales (Swanson, Maas y Buckwalter, 1994).

La relación entre una variable independiente y una dependiente se puede entender mejor si trazamos dos ejes perpendiculares uno del otro. Uno representa a la variable independiente; el otro, a la dependiente. (Cuando dos ejes forman ángulos rectos entre sí se denominan ejes *ortogonales*.) De acuerdo a la tradición matemática, x, la variable independiente, es el eje horizontal y y, la dependiente, representa el eje vertical (x se denomina la abscisa y y la ordenada). Los valores para x se grafican en el eje de las x, y los valores de y en el eje de las y.

Una forma común y útil de "ver" e interpretar una relación es graficar un par de valores de xy, usando los ejes x y y como marco de referencia. En un estudio de desarrollo infantil, supongamos que tenemos dos grupos de medidas. Las medidas x de edad cronológica y las medidas y representan edad lectora. *La edad lectora* se denomina también edad de crecimiento. Las mediciones en serie de diferentes áreas del crecimiento de los individuos —por ejemplo estatura, peso o inteligencia— se expresan como la edad cronológica promedio en la que aparecen en la población estándar.

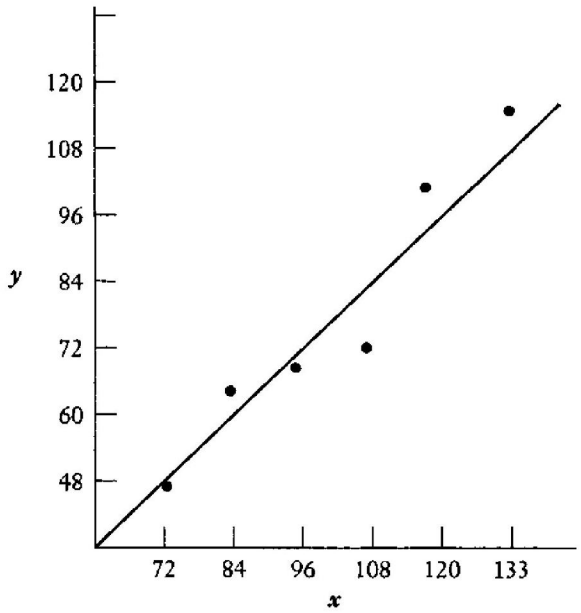
x: edad cronológica (en meses)	y: edad lectora (en meses)
72	48
84	62
96	69
108	71
120	100
132	112

Estas medidas se grafican en la figura 3.3.

La relación entre edad cronológica (EC) y edad lectora (EL), ahora puede "verse" y aproximarse de forma burda. Observe que hay una tendencia pronunciada (como se podría esperar), para que una mayor EC se asocie con una mayor EL, una EC media con una EL media, y una EC menor, con una EL menor. En otras palabras, la relación entre las variables independiente y dependiente, en este caso EC y EL, puede observarse en una gráfica como la que aparece en la figura 3.3. Se ha trazado una recta para "mostrar" esta relación: constituye un promedio aproximado de todos los puntos de la gráfica. Observe que si uno conoce las medidas de la variable independiente y una relación como la que se muestra en la figura 3.3, uno puede predecir, con considerable precisión, las medidas de la variable dependiente. Gráficas como ésta pueden usarse, desde luego, con cualesquier grupo de medidas para variables dependientes e independientes.

El estudiante debe estar atento a la posibilidad de que en un estudio una variable sea independiente, mientras en otro sea dependiente, e inclusive ambas en un mismo estudio.

FIGURA 3.3



Un ejemplo es la satisfacción laboral. La mayoría de los estudios sobre satisfacción laboral la utilizan como variable dependiente. Day y Schoenrade (1997) muestran el efecto de la orientación sexual en las actitudes laborales. Una de estas actitudes laborales es la satisfacción laboral. De la misma forma, Lekewise, Hodson (1989) estudia las diferencias de género en la satisfacción laboral. Scott, Moore y Miceli (1997) encuentran a la satisfacción laboral ligada a los patrones de comportamiento de los adictos al trabajo. Hay estudios en donde la satisfacción laboral es usada como una variable independiente: Meiksins y Watson (1989) muestran cuánto influye la satisfacción laboral en la autonomía profesional de los ingenieros. Estudios de Somers (1996); Francis-Felsen, Coward, Hogan y Duncan (1996); y Hutchinson y Turner (1988) evaluaron el efecto de la satisfacción laboral en la rotación del personal de enfermería.

Otro ejemplo es la ansiedad, que se ha estudiado como una variable independiente que afecta a la variable dependiente aprovechamiento. Oldani (1997) encontró que la ansiedad de la madre durante el embarazo influye en el aprovechamiento de los hijos (medido como éxito en el área musical). Capaldi, Crosby y Stoolmiller (1996) emplearon los niveles de ansiedad de varones adolescentes para predecir el momento de su primer encuentro sexual. Onwuegbuzie y Seaman (1995) estudiaron los efectos de la ansiedad ante los exámenes en la realización de la prueba, en un curso de estadística. La ansiedad también puede concebirse y usarse como una variable dependiente: por ejemplo, puede utilizarse para estudiar la diferencia entre tipos de cultura, nivel socioeconómico y género (véase Guida y Ludlow, 1989; Murphy, Olivier, Monson y Sobol, 1991). En otras palabras, la clasificación de la variable independiente y la dependiente es en realidad una taxonomía de los usos de la variable más que una distinción entre diferentes tipos de variables.

Variables activas y variables atributo

Una clasificación que nos será útil en nuestro estudio posterior del diseño de la investigación se basa en la distinción entre variables experimentales y medidas. Cuando se planea y ejecuta la investigación es importante distinguir entre estos dos tipos de variables. Las variables manipuladas se llamarán variables *activas*, mientras que las variables medidas se denominarán variables *atributo*. Por ejemplo, Colwell, Foreman y Trotter (1993) compararon dos métodos de tratamiento de las úlceras de presión de los pacientes encamados. Las variables dependientes fueron eficacia y efectividad costo. Los dos métodos de tratamiento fueron una compresa de gasa humedecida y una compresa con una cubierta de hidrocoloide. Los investigadores controlaron quién recibía qué tipo de tratamiento. Como tal, el tratamiento o variable independiente fue una variable activa o manipulada.

Así, cualquier variable manipulada, constituye una variable activa. "Manipulación" significa, en esencia, hacer cosas diferentes a distintos grupos de sujetos, como se verá con claridad en un capítulo posterior al discutir a profundidad las diferencias entre investigación experimental y no experimental. Se dice que existe manipulación cuando un investigador hace algo a un grupo (por ejemplo, reforzar positivamente cierta clase de comportamiento) y hace algo distinto con el otro grupo, o tiene a dos grupos siguiendo diferentes instrucciones. Cuando uno usa diferentes métodos de enseñanza o premia a los sujetos de un grupo y castiga a los de otro, o crea ansiedad a través de instrucciones que generan preocupación, uno está *activamente* manipulando las variables: métodos, reforzamiento y ansiedad.

Otra clasificación relacionada, usada principalmente por los psicólogos es la de las variables *estímulo* y *respuesta*. Una *variable estímulo* es cualquier condición o manipulación del ambiente realizada por el experimentador, que evoque una respuesta en un organismo. Una *variable de respuesta* es cualquier clase de comportamiento del organismo. El supuesto es que para cualquier clase de comportamiento siempre hay un estímulo. Por lo tanto, el comportamiento del organismo es una respuesta. Esta clasificación se refleja en la bien conocida ecuación: $R = f(O, E)$, que se lee: "las respuestas son una función del organismo y los estímulos", o "las variables de respuesta son una función de las variables organísmicas y de las variables estímulo".

Las variables que no pueden ser manipuladas son *las atributo o características del sujeto*. Es imposible, o al menos muy difícil, manipular muchas variables. Las variables consistentes en características humanas como inteligencia, aptitud, género, estatus socioeconómico, conservadurismo, dependencia de campo, necesidad de logro y actitudes son variables atributo. Los sujetos llegan a nuestro estudio con estas variables (atributos) ya presentes o preexistentes. El entorno temprano, la herencia, y otras circunstancias han hecho de los individuos lo que son. Se les llama también variables *organísmicas*. Cualquier propiedad, característica o atributo de un individuo constituye una variable organísmica, digamos que forma parte del organismo. En otras palabras, las variables organísmicas son aquellas características que los individuos poseen en diversos grados cuando ingresan a la situación de investigación. El término *diferencias individuales* implica variables organísmicas. Una de las más comunes variables atributo en las ciencias sociales y del comportamiento es el género: femenino-masculino. Los estudios diseñados para comparar diferencias de género involucran una variable atributo. Tomemos, por ejemplo, el estudio de Weerth y Kalma (1993). Estos investigadores compararon a hombres y mujeres en su respuesta a la infidelidad del cónyuge o pareja. La variable atributo aquí es género, que *no* es una variable manipulada. Hay estudios donde las puntuaciones de una o varias pruebas se usaron para dividir a un conjunto de personas en dos o más grupos. En este caso, las diferencias del grupo se reflejan como una variable atributo, como lo ilustra el estudio de Hart, Forth y

Hare (1990) quienes aplicaron una prueba de psicopatología a varones recluidos en una prisión. Con base en sus puntuaciones, los internos fueron asignados a uno de tres grupos: bajo, medio y alto. Después se comparó su puntuación en una batería de pruebas neuropsicológicas. El nivel de psicopatología preexiste y el investigador no lo manipula. Si un interno puntuaba alto, era asignado al grupo alto. Así, la psicopatología es una variable atributo en este estudio. Hay algunos estudios donde la variable independiente podría haber sido manipulada; sin embargo, por razones logísticas o legales no lo fue. Un ejemplo es el estudio de Swanson, Maas y Buckwalter (1994) quienes compararon diferentes formas de atención y su efecto en las medidas cognitivas y funcionales de los pacientes con Alzheimer. La variable atributo fue el tipo de atención. No se permitió a los investigadores asignar a sus pacientes a dos diferentes instituciones de atención (tradicional *vs.* unidad de cuidado especial). Los investigadores se vieron forzados a estudiar a los sujetos una vez que habían sido asignados al centro correspondiente. Por ello puede considerarse que la variable independiente es una variable no manipulada. Los investigadores heredaron grupos intactos.

La palabra ***atributo*** es lo suficientemente precisa cuando se usa con objetos o referentes inanimados. Sin embargo, las organizaciones, instituciones, grupos, poblaciones, casas y áreas geográficas también poseen atributos —*atributos activos*—. Las organizaciones son productivas de manera variable; las instituciones pasan de moda; los grupos difieren en su cohesión; las áreas geográficas varían ampliamente en sus recursos.

Esta distinción de atributo activo es general, flexible y útil. Veremos que algunas variables, por su propia naturaleza, son siempre atributos, mientras que otras variables que son atributos pueden también ser activas. Esta última característica hace posible investigar las "mismas" relaciones de diferentes formas. Y usando de nuevo el ejemplo de la variable ansiedad, podemos medirla: como es evidente en este caso, es una variable atributiva. Sin embargo, también puede ser manipulada al inducir diferentes grados de ansiedad: por ejemplo, si les decimos a los sujetos de un grupo experimental que la tarea que van a realizar va a ser muy difícil, que su inteligencia será evaluada y que su futuro depende de la puntuación que obtengan. A los sujetos del otro grupo experimental les decimos que lo hagan lo mejor posible, pero relajados, y que el resultado no es importante y que no va a influir en su futuro. En realidad no podemos asumir que la ansiedad medida (atributo) y la ansiedad manipulada (activa) sean la misma. Podemos suponer que ambas son "ansiedad" en un sentido amplio, pero ciertamente no son iguales.

Variables continuas y categóricas

Ya se ha hecho una distinción entre variables continuas y categóricas en especial útil para la planeación de la investigación y el análisis de datos. Sin embargo, su importancia justifica una consideración más amplia.

Una variable *continua* es capaz de asumir un conjunto ordenado de valores dentro de cierto rango. Esta definición significa, primero, que el valor de una variable continua refleja al menos un orden categórico, que un mayor valor de la variable implica más de la propiedad en cuestión que un menor valor. Los valores producidos por una escala para medir dependencia, por ejemplo, expresan diferentes cantidades de dependencia, desde la alta pasando por la media hasta la baja. Segundo, las medidas continuas en uso están contenidas en un rango, y cada individuo obtiene una "puntuación" dentro del mismo. Una escala para medir dependencia puede tener un rango de uno a siete. La mayoría de las escalas usadas en ciencias del comportamiento también tienen una tercera característica: hay un conjunto teóricamente infinito de valores en el rango. (Las escalas de rangos

ordenados son algo diferentes; se analizarán más adelante en el libro.) Esto quiere decir que una puntuación de un individuo en particular puede ser de 4.72 más que sólo de 4 o 5.

Las variables *categóricas*, como las llamaremos, pertenecen a una clase de mediciones llamadas nominales (se explicarán en el capítulo 25). En una medición nominal hay dos o más subconjuntos del grupo de objetos que se mide. Se categoriza a los individuos en razón de la posesión de las características que definen cualquier subgrupo. "Categorizar" significa asignar un objeto a una subclase (o subconjunto) de una clase (o conjunto) con base en que el objeto posea o no la característica que define al subconjunto. El individuo que está en proceso de ser categorizado posee o no la propiedad definitoria; esto es un asunto de todo o nada. Los ejemplos más simples son las variables categóricas dicotómicas: femenino-masculino, republicano-demócrata, correcto-incorrecto. Las variables politómicas, que son aquellas con más de dos subconjuntos, son bastante comunes, en especial en sociología y economía: preferencia religiosa, nivel educativo, nacionalidad y elección laboral, entre otras.

Las variables categóricas y la medición nominal tienen requisitos simples: se considera iguales a todos los miembros de un subconjunto y todos tienen asignado el mismo nombre (nominal) y el mismo número. Si la variable es preferencia religiosa, por ejemplo, todos los protestantes son iguales, todos los católicos son iguales y todos los "otros" son iguales. Si un individuo es católico (definido operacionalmente de una manera apropiada), la persona es asignada a la categoría "católica" y también se le asigna un "1" en esa categoría. Es decir, se contabiliza a ese individuo como "católico". Las variables categóricas son "democráticas". No hay un orden en cuanto a rango, ni mayor que o menor que, entre las categorías y se asigna un mismo valor a todos los miembros de una categoría.

La expresión "variables cualitativas" algunas veces se ha aplicado a las variables categóricas, en particular a las dicotómicas, probablemente en contraste con las "variables cuantitativas" (nuestras variables continuas). Este uso refleja un concepto distorsionado de lo que son las variables, ya que siempre son cuantificables; si no lo fueran, no serían variables. Si x tiene solamente dos subconjuntos y puede asumir sólo dos valores (1 y 0), éstos siguen siendo valores, y la variable varía. Si x es una variable politómica, como la afiliación política, cuantificamos nuevamente al asignar valores enteros a los individuos. Si un individuo dice ser demócrata, se ubica a esa persona en el subgrupo demócrata y se le asigna un 1. Todos los individuos en el subconjunto demócrata tendrán un valor de 1. Es en extremo importante comprender esto porque, por principio, son las bases para cuantificar muchas variables, incluso tratamientos experimentales, para llevar a cabo análisis complejos. En el análisis de regresión múltiple, como veremos más adelante, todas las variables —continuas y categóricas— son ingresadas como variables al análisis. En el ejemplo de género que se mencionó arriba, 1 era asignado a un género y 0 al otro. Diseñamos una columna de unos y ceros de la misma forma en que estableceríamos una columna de puntuaciones de dependencia. La columna de unos y ceros representa la cuantificación de la variable género. Aquí no hay misterio. Estas variables han sido llamadas variables *prototipo* (conocidas como dummy en inglés). Como son muy útiles y poderosas e incluso indispensables en el análisis de datos de la investigación moderna, requieren ser entendidas con claridad. Una explicación más a fondo puede encontrarse en Kerlinger y Pedhazur (1973) y en el capítulo 34 de este libro. El método se aplica con facilidad a las politomías. Una *politomía* es una división de los miembros de un grupo en tres o más subdivisiones.

Constructos observables y variables latentes

En gran parte de la discusión previa de este capítulo se ha implicado —pues no se ha declarado de forma explícita— que existe una diferencia fundamental entre constructos y

variables observadas. Más aún, podemos afirmar que los constructos no son observables; y que las variables, cuando se definen operacionalmente, son observables. Esta distinción es importante en tanto que si no estamos plenamente conscientes del nivel de discurso en que nos encontramos al hablar acerca de variables, es difícil ser claros sobre lo que hacemos.

Una expresión fructífera e importante que se encontrará y usará extensamente en este libro es "variable latente". Una variable latente es una "entidad" no observada, que se presume subyace a las variables observadas. El ejemplo mejor conocido de una variable latente importante es "inteligencia". Podemos decir que tres pruebas de habilidad —verbal, numérica y espacial— están relacionadas de manera positiva y sustancial. Esto significa, en general que las personas con puntuación alta en una, tienden a tener altas puntuaciones en las otras; de forma similar quienes obtienen bajas puntuaciones en una, tenderán a presentar bajas en las otras. Creemos que hay algo común a las tres pruebas o variables observadas, y lo llamamos "inteligencia", que es una variable latente.

Hemos encontrado muchos ejemplos de variables latentes en las páginas previas: aprovechamiento, creatividad, clase social, satisfacción laboral, preferencia religiosa, etcétera. De hecho, siempre que mencionamos los nombres de fenómenos en que varían las personas o los objetos, hablamos de variables latentes. En el campo de la ciencia, nuestro interés real está más en las relaciones entre variables latentes que entre variables observadas, ya que buscamos explicar fenómenos y sus relaciones. Cuando enunciamos una teoría, enunciamos en parte relaciones sistemáticas entre variables latentes. No estamos demasiado interesados en la relación entre el comportamiento observado de frustración y la conducta observada de tipo agresivo, por ejemplo, aunque debemos trabajar con ellos en el nivel empírico. En realidad, estamos interesados en la relación entre la variable latente frustración y la variable latente agresión.

Debemos ser precavidos, sin embargo, cuando tratamos con variables no observables. Los científicos que usan términos como "hostilidad", "ansiedad" y "aprendizaje", están conscientes de que hablan a cerca de constructos inventados. La "realidad" de estos constructos se infiere a partir del comportamiento. Si desean estudiar diferentes tipos de motivación, deben saber que "motivación" es una variable latente, un constructo inventado para dar cuenta de un comportamiento presumiblemente "motivado". Es necesario que sepan que esta "realidad" sólo está postulada. Sólo pueden juzgar si los jóvenes están motivados o no al observar sus comportamientos. Aun así para estudiar la motivación, deben medirla o manipularla. Pero no pueden medirla de forma directa por que, en pocas palabras, es una variable que está "en la cabeza", una entidad no observable, una variable latente. Se inventó el constructo "por algo" que *se presume que está* dentro de los individuos, "algo" que los impulsa a comportarse de tal y cual manera. Esto significa que los investigadores deben medir siempre supuestos indicadores de motivación y no a ella en sí misma. En otras palabras, deben medir algún tipo de comportamiento, sean marcas en un papel, palabras habladas, o gestos significativos, y después hacer inferencias sobre características supuestas o variables latentes.

Se han usado otros términos para expresar más o menos las mismas ideas. Por ejemplo, Tolman (1951 pp. 115-129) llamó a los constructos variables intervinientes. Las *variables intervinientes* representan un término inventado para dar cuenta de procesos psicológicos no observables, internos, que a su vez dan cuenta de la conducta. Una variable interviniente es una variable "en la cabeza": No se le puede ver, oír o tocar. Se infiere a partir del comportamiento. La "hostilidad" se infiere de actos presumiblemente hostiles o agresivos. La "ansiedad" se infiere de la puntuación en una prueba, respuesta en la piel, frecuencia cardiaca y ciertas manipulaciones experimentales. Otro término es "constructo hipotético", cuyo significado es bastante similar al de variable latente, aunque con un poco menos de generalidad; no es necesario detenernos en él. Debemos mencionar, sin embar-

go, que "variable latente" parece ser una expresión más general y aplicable que "variable interviniente" y que "constructo hipotético", en tanto que puede usarse virtualmente para cualquier fenómeno que se presume influye o es influido por otro. En otras palabras "variable latente" puede utilizarse con fenómenos psicológicos, sociológicos y de otro tipo. "Variable latente" parece ser un vocablo más útil por su generalidad y también por que ahora es posible, en el enfoque de análisis de estructuras de covarianza, evaluar el efecto de las variables latentes entre sí y las llamadas variables manifiestas u observadas. Está discusión que parece muy abstracta después se concretará para ser, esperamos, significativa. Veremos entonces que la idea de las variables latentes y las relaciones entre ellas son extremo importantes, fructíferas y útiles, y ayudan a cambiar los enfoques fundamentales para afrontar los problemas de investigación.

Ejemplos de variables y definiciones operacionales

Hemos aportado una cantidad de constructos y definiciones operacionales. Para ilustrar y aclarar la discusión previa, en especial en lo referente a la diferencia entre variables experimentales y variables medidas, y entre constructos y variables definidas operacionalmente, presentamos diversos ejemplos. Si la definición es experimental se etiqueta con (E); si es una definición medida se señala como (M).

Las definiciones operacionales difieren en su grado de especificidad. Algunas están ligadas de manera estrecha a las observaciones. Definiciones de "pruebas", como "la inteligencia se define como una puntuación x en una prueba de inteligencia" son muy específicas. Una definición como "la frustración consiste en no alcanzar una meta" son más generales y requieren una mayor especificación para que sean medibles.

Clase social. "...dos o más grupos de personas que se cree que están en una posición social superior e inferior y que son así categorizados por los miembros de una comunidad" (M) (Warner y Lunt, 1941, p. 82). Para ser operacional, está definición ha de especificarse a partir de preguntas dirigidas a las creencias de la gente sobre las posiciones de otras personas. Se trata de una definición subjetiva de clase social. La clase social, o el estatus social, también es define de forma más objetiva a través de índices como ocupación, ingreso y educación, o por combinaciones de tales índices. Por ejemplo, "...transformamos la información acerca de educación, ocupación e ingresos de los padres de jóvenes de una liga juvenil en un índice de nivel socioeconómico (NSE), en el que las puntuaciones altas indican una avanzada educación, una ocupación prestigiosa y un ingreso desahogado. Calificaciones menores reflejan pobreza, educación incompleta y los trabajos más modestos" (M) (Herrnstein y Murray, 1996, p. 131).

Aprovechamiento (escolar, aritmético y en ortografía). El aprovechamiento se acostumbra definir operacionalmente a partir de una prueba estandarizada de aprovechamiento (por ejemplo la prueba de Iowa de Destrezas Básicas, la prueba de aprovechamiento o la prueba elemental de la batería de Evaluación de Infantil de Kaufman [K-ABC]), por medio del promedio o por el juicio del maestro. "El aprovechamiento del estudiante se midió por una puntuación combinada de pruebas de lectura y matemáticas" (M) (Peng y Wright, 1994). En ocasiones, el aprovechamiento se presenta en forma de una prueba de ejecución. Silverman (1993) examinó en un grupo de estudiantes dos habilidades del juego de voleibol: la prueba de servicio y la prueba del pase con antebrazo. En la primera, los estudiantes recibían una puntuación entre 0 y 4, en función de donde fuera colocado el balón servido. La prueba del pase con antebrazo consistía en hacer rebotar la pelota en el antebrazo. El criterio usado fue contar el número de veces en que un estudiante pudo pasar el balón más allá de una línea de dos metros y medio contra la pared en un periodo de un minuto (M). En algunos estudios educativos también se usa una definición operacional del

concepto *percepción del aprovechamiento del estudiante*. Aquí se pide a los estudiantes que se evalúen a sí mismos. La pregunta usada por Shoffner (1990) fue: "¿Qué clase de estudiante crees que eres?". Las opciones eran "estudiante de 9 y 10" "estudiante de 8" y "estudiante de 7 y 6" (M).

Aprovechamiento (desempeño académico). "Como resultado se obtuvieron las calificaciones de todos los estudiantes en todas las secciones y se usaron para determinar la categoría de cada estudiante participante en el estudio. Se calculó el rango porcentilar por sección para cada uno y se usó como la medida dependiente del aprovechamiento en el análisis final de los datos" (M) (Strom, Hocevar y Zimmer, 1990).

Motivación intrínseca se define operacionalmente por Hom, Berger *et al.* (1994) como "la cantidad acumulada de tiempo que cada estudiante juega con un patrón de bloques sin un sistema de reforzamiento" (M).

Popularidad. La popularidad con frecuencia se define operacionalmente por el número de elecciones sociométricas que un individuo recibe de otros (en su clase, grupo de juego, etcétera). Se le pregunta a los sujetos: "¿con quién te gustaría trabajar?", "¿con quién te gustaría jugar?" y otras cuestiones similares. Cada sujeto debe elegir a uno, dos o más individuos de su grupo con base en esas preguntas de criterios (M).

Compromiso con el trabajo "...el comportamiento de cada niño durante una lección fue evaluada cada seis segundos como apropiadamente comprometida o no. La puntuación del compromiso con el trabajo, por lección, fue el porcentaje de las unidades de seis segundos en que los niños fueron calificados como apropiadamente comprometidos" (M) (Kounin y Doyle, 1975).

Reforzamiento. Las definiciones de reforzamiento tienen diversas formas. La mayoría incluye, de una forma u otra, el principio de la recompensa. Sin embargo, se puede utilizar tanto el reforzamiento positivo como negativo. A continuación se presentan definiciones experimentales específicas.

En los siguientes 10 minutos cada opinión que el sujeto (S) enunció fue registrada por el experimentador (E) y reforzada. Para dos grupos, el E estuvo de acuerdo con la opinión enunciada al decir "sí tienes razón", "así es", o algo similar, o al sentir o sonreír en caso de que no pudiera interrumpir (E).

> ...se administraron al modelo y al niño de manera alternativa doce diferentes grupos de reactivos de historias... Ante cada uno de ellos, el modelo expresaba de forma consistente respuestas en forma de juicios opuestas a la orientación moral del niño... Y el experimentador reforzaba el comportamiento del modelo con respuestas de aprobación verbal tales como "muy bien", "está bien" y "qué bueno". El niño fue reforzado de manera parecida cada vez que adoptaba el tipo de juicios morales del modelo en respuesta a su propio grupo de reactivos [esto se denomina "reforzamiento social"] (E) (Bandura y MacDonald, 1994).

El maestro otorga al niño un reconocimiento verbal cada vez que exhibe el comportamiento deseado. Éstos son: atender a la instrucción, cumplir con el trabajo escolar y contestar en voz alta. El registro se hace cada 15 segundos (E) (Martens, Hiralall y Bradley, 1997).

Actitudes hacia el SIDA se define en una escala de 18 reactivos, cada uno con un formato de tipo Likert para reflejar diversas actitudes hacia los pacientes con SIDA. Algunos ejemplos de los reactivos son: "a la gente con SIDA no se le debería permitir usar los sanitarios públicos", y "debería ser obligatorio para todos los estadounidenses hacerse una prueba para SIDA" (M) (Lester, 1989).

Personalidad limítrofe. Comrey (1993) la define como la presencia de una baja puntuación en tres escalas de la Escala de Personalidad de Comrey: confianza *vs.* defensividad, conformidad social *vs.* rebeldía, y estabilidad emocional *vs.* neuroticismo.

Delincuencia laboral se define operacionalmente como una combinación de tres variables: número de accidentes imputables al sujeto, número de cartas de advertencia y número de suspensiones (M) (Hogan y Hogan, 1989).

Religiosidad se define como una puntuación en la Escala de Francis de Actitudes hacia la Cristiandad que consta de 24 reactivos con una escala de respuesta tipo Likert. Algunos reativos son: "Orar me ayuda mucho" y "Dios me guía para conducir una vida mejor" (M) (Gillings y Joseph, 1996). La religiosidad no debe confundirse con la preferencia religiosa. La religiosidad se refiere a la fuerza de la devoción a la religión que uno ha elegido.

Autoestima es una variable independiente manipulada en el estudio de Steele, Spencer y Lynch (1993). En este caso, se aplica a los sujetos una prueba de autoestima, pero cuando se les retroalimenta la información en el reporte de retroalimentación, que parece ser el oficial es ambigua. Se divide a los sujetos del mismo nivel de autoestima medida en tres grupos diferentes de retroalimentación: positiva, negativa y ausente. En la condición de retroalimentación positiva (autoestima positiva), se describe a los sujetos con enunciados tales como "pensamiento claro". Aquéllos en el grupo negativo (autoestima negativa) reciben adjetivos como "pasivos al actuar". Al grupo "sin retroalimentación" se les indica que su perfil de personalidad (autoestima) no estuvo listo por demoras en la calificación e interpretación (E). La mayoría de los estudios sobre autoestima usan una definición operacional medida. En el ejemplo anterior, Steele, Spencer y Lynch también usaron la Escala de Sentimientos de Inadecuación de Autoestima de Janis-Field (M). En otro ejemplo Luhtanen y Crocker (1992) definieron la autoestima colectiva como una puntuación en una escala de 16 reactivos tipo Likert en los que se solicitaba pensar a cerca de una variedad de grupos sociales y características de sus miembros, tales como género, religión, raza y grupo étnico (M).

La *raza* es por lo general una variable medida. Sin embargo en un estudio de Annis y Corenblum (1986), un experimentador (E) ya sea de raza blanca o india preguntó a 83 niños indios canadienses de nivel preescolar y primer año sobre preferencias raciales e identidad personal. El interés se centraba en averiguar si la raza del experimentador influía o no en las respuestas.

Soledad. Una definición de ésta es la puntuación en la Escala de Soledad de UCLA que incluye reactivos tales como: "nadie me conoce realmente bien" o "carezco de compañía". También se cuenta con la Escala de Deprivación y Soledad que presenta reactivos como: "experimento una sensación de vacío" o "no hay quien muestre un interés particular en mí" (M) (Oshagan y Allen, 1992).

Halo. Se han postulado muchas definiciones operacionales del efecto de halo. Balzer y Sulsky (1992) encontraron y resumieron 108 definiciones que ajustaron en 6 categorías. Una establece que halo es "...la variaza promedio intratasa o la desviación estándar de las evaluaciones". Otra puede ser: "Comparar las evaluaciones obtenidas con las proporcionadas por jueces expertos" (M).

Memoria: recuerdo y reconocimiento "...recuerdo significa pedir al participante que repita lo que recuerda de los reactivos que le fueron mostrados, y asignar un punto por cada uno que corresponda a la lista de estímulos inicial"(M) (Norman, 1976, p. 97). "La prueba de reconocimiento consistió en 62 frases presentadas a todos lo sujetos... que fueron instruidos para evaluar cada frase de acuerdo a su grado de confianza de que la oración hubiera sido presentada en el conjunto inicial" (M) (Richter y Seay, 1987).

Habilidades sociales. Pueden ser definidas operacionalmente como una puntuación en la Escala de Evaluación de Destrezas Sociales (Gresham y Elliot, 1990). Existe la posibilidad de contar con información del estudiante de su padre y del maestro. Se evalúan los comportamientos sociales en términos de frecuencia de ocurrencia y también de acuerdo a su nivel de importancia. Algunos reactivos incluyen: "se lleva bien con gente que es

diferente (maestro)", "se ofrece a ayudar a miembros de la familia con sus tareas (papá)", y, "cuestiono cortésmente las reglas que pueden ser injustas (estudiante)" (M).

Congraciamiento. Una de las muchas técnicas de manejo de impresión (véase Orpen, 1996; Gordon, 1996). Se define operacionalmente como una puntuación en la Escala de Kumar y Beyerlein (1991). Que comprende 25 reactivos tipo Likert diseñados para medir la frecuencia en que el subordinado, en una relación superior-subordinado usa tácticas para congraciarse (M). Strutton, Pelton y Lumpkin (1995) modificaron la Escala de Kumar-Beyerlein, para medir el congraciamiento entre el vendedor y cliente (M).

Feminismo. Se define por la puntuación en el Cuestionario de Actitudes hacia las mujeres. Este instrumento consta de 18 enunciados en los que el entrevistado registra su acuerdo en una escala de cinco puntos. Los reactivos incluyen: "los hombres han mantenido el poder por demasiado tiempo"; "los concursos de belleza son degradantes para la mujer"; "los niños de madres trabajadoras tienden a sufrir" (Wilson y Reading, 1989).

Valores. "Ordene las 10 metas de acuerdo a la importancia que tienen para usted: 1) Tener éxito financiero; 2) Ser querido; 3) Tener éxito en el ámbito familiar; 4) Ser capaz en lo intelectual; 5) Vivir de acuerdo a principios religiosos; 6) Ayudar a otros; 7) Ser normal, bien ajustado; 8) Cooperar con los demás; 9) Trabajar con detalle; 10) Alcanzar el éxito laboral" (M) (Newcomb, 1978).

Democracia (democracia política). "El índice (de democracia política) consiste en tres indicadores de soberanía popular y tres de libertades políticas. Las medidas de soberanía popular son: 1) Elecciones limpias, 2) Selección ejecutiva efectiva, y 3) Selección legislativa. Los indicadores de libertades políticas son: 4) Libertad de prensa, 5) Libertad para que la oposición se asocie y 6) Sanciones gubernamentales" (M). Bollen (1979) proporciona detalles operacionales de seis indicadores sociales en un apéndice (pp. 585-586). Éste es un ejemplo particularmente bueno de la definición operacional de un concepto complejo. Más aún, constituye una excelente descripción de los ingredientes de la democracia.

Los beneficios del pensamiento operacional han sido grandiosos. De hecho el operacionalismo ha sido y es uno de los movimientos más significativos e importantes de nuestro tiempo. El operacionalismo extremo, por supuesto, puede ser peligroso porque oscurece el reconocimiento de la importancia de los constructos y definiciones constitutivas en la ciencia del comportamiento y porque puede restringir la investigación a problemas triviales. Sin embargo, hay poca duda de que sea una sana influencia. Resulta una clave indispensable para alcanzar la objetividad (sin la cual no hay ciencia), en tanto que demanda que las observaciones sean públicas y replicables, lo que ayuda a colocar las actividades de investigación más allá de los investigadores y sus predilecciones. Y, como dijo Underwood (1957, p. 53) en su texto clásico sobre investigación psicológica:

> Yo diría que el pensamiento operacional hace mejores científicos. El operacionalista se ve forzado a sacudir y aclarar sus conceptos empíricos... el operacionalismo facilita la comunicación entre científicos ya que el significado de los conceptos así definidos no es sujeto fácilmente a una mala interpretación.

RESUMEN DEL CAPÍTULO

1. Un *concepto* es una expresión de una abstracción formada a partir de la generalización de un particular, por ejemplo, peso. Esta expresión se deriva de observaciones de ciertos comportamientos o acciones.
2. Un *constructo* es un concepto que se ha formulado para ser usado en la ciencia. Se usa en esquemas teóricos y se define de tal manera que sea susceptible de ser observado y medido.

¿Qué son las hipótesis?

Hipótesis Explicaciones tentativas del fenómeno investigado que se enuncian como proposiciones o afirmaciones.

Las **hipótesis** son las guías de una investigación o estudio.[1] Las hipótesis indican lo que tratamos de probar y se definen como explicaciones tentativas del fenómeno investigado. Se derivan de la teoría existente y deben formularse a manera de proposiciones. De hecho, son respuestas provisionales a las preguntas de investigación. Cabe señalar que en nuestra vida cotidiana constantemente elaboramos hipótesis acerca de muchas cosas y luego indagamos su veracidad. Por ejemplo, establecemos una pregunta de investigación: "¿Le gustaré a Paola?", y una hipótesis: "Le resulto atractivo a Paola". Esta hipótesis es una explicación tentativa y está formulada como proposición. Después investigamos si se acepta o se rechaza la hipótesis, al cortejar a Paola y observar el resultado.

1

¿En toda investigación cuantitativa debemos plantear hipótesis?

No, no en todas las investigaciones cuantitativas se plantean hipótesis. El hecho de que formulemos o no hipótesis depende de un factor esencial: el alcance inicial del estudio. Las investigaciones cuantitativas que formulan hipótesis son aquellas cuyo planteamiento define que su alcance será correlacional o explicativo, o las que tienen un alcance descriptivo, pero que intentan pronosticar una cifra o un hecho. Esto se resume en la tabla 6.1.

Tabla 6.1 Formulación de hipótesis en estudios cuantitativos con diferentes alcances

Alcance del estudio	Formulación de hipótesis
Exploratorio	No se formulan hipótesis.
Descriptivo	Sólo se formulan hipótesis cuando se pronostica un hecho o dato.
Correlacional	Se formulan hipótesis correlacionales.
Explicativo	Se formulan hipótesis causales.

Un ejemplo de estudio con alcance descriptivo y pronóstico sería aquel que únicamente pretenda medir el índice delictivo en una ciudad (no se busca relacionar la incidencia delictiva con otros factores como el crecimiento poblacional, el aumento de los niveles de pobreza o la drogadicción, ni mucho menos establecer las causas de tal índice). Entonces, tentativamente se pronosticaría, mediante una hipótesis, cierta cifra o proporción: "el índice delictivo para el siguiente semestre será menor a un delito por cada mil habitantes".

Por lo regular, los estudios cualitativos no formulan hipótesis antes de recolectar datos (aunque no siempre ocurre así). Su naturaleza es más bien inducir las hipótesis por medio de la recolección y el análisis de los datos, como se comentará en la tercera parte del libro, "El proceso de la investigación cualitativa".

En una investigación podemos tener una, dos o varias hipótesis.

¿Las hipótesis son siempre verdaderas?

Las hipótesis no necesariamente son verdaderas, pueden o no serlo, y pueden o no comprobarse con datos. Son explicaciones tentativas, no los hechos en sí. Al formularlas, el investigador no está totalmente seguro de que vayan a comprobarse. Como ejemplifican Black y Champion (1976), una hipó-

[1] El término proviene del verbo griego *hypotithenai*, que significa "suponer" o "colocar por debajo" (Williams, 2003).

tesis es diferente de la afirmación de un hecho. Si alguien establece la siguiente hipótesis (refiriéndose a un país determinado): "las familias que viven en zonas urbanas tienen menor número de hijos que las familias que viven en zonas rurales", ésta puede ser o no comprobada. En cambio, si una persona sostiene lo anterior basándose en información de un censo poblacional reciente de ese país, no establece una hipótesis, sino que afirma un hecho.

En el ámbito de la investigación científica, las hipótesis son proposiciones tentativas acerca de las relaciones entre dos o más variables y se apoyan en conocimientos organizados y sistematizados. Una vez que se prueba una hipótesis, tiene un impacto en el conocimiento disponible que puede modificarse y por consiguiente, pueden surgir nuevas hipótesis (Davis, 2008 e Iversen, 2003).

Las hipótesis pueden ser más o menos generales o precisas, y abarcar dos o más variables; pero en cualquier caso son sólo afirmaciones sujetas a comprobación empírica, es decir, a verificación en la realidad.

Ejemplo

Hipótesis

- "La proximidad geográfica entre los hogares de las parejas de novios está vinculada positivamente con el nivel de satisfacción que les proporciona su relación."
- "La incidencia de cáncer pulmonar es mayor entre los fumadores que entre los no fumadores."
- "A mayor variedad en el trabajo, habrá mayor motivación intrínseca para cumplirlo."

Observe que, por ejemplo, la primera hipótesis vincula dos variables: "proximidad geográfica entre los hogares de los novios" y "nivel de satisfacción en la relación".

¿Qué son las variables?

En este punto es necesario definir qué es una variable. Una **variable** es una propiedad que puede fluctuar y cuya variación es susceptible de medirse u observarse.[2] Ejemplos de variables son el género, la presión arterial, el atractivo físico, el aprendizaje de conceptos, la religión, la resistencia de un material, la masa, la personalidad autoritaria, la cultura fiscal y la exposición a una campaña de propaganda política. El concepto de variable se aplica a personas u otros seres vivos, objetos, hechos y fenómenos, los cuales adquieren diversos valores respecto de la variable referida. Por ejemplo, la inteligencia, ya que es posible clasificar a las personas de acuerdo con su inteligencia; no todas las personas la poseen en el mismo nivel, es decir, varían en inteligencia.

Variable Propiedad que tiene una variación que puede medirse u observarse.

1

Otros ejemplos de variables son: el rendimiento de cierta especie de semilla, la eficacia de un procedimiento de construcción, el tiempo que tarda en manifestarse una enfermedad y otros. En todos los casos se producen variaciones.

Las variables adquieren valor para la investigación científica cuando llegan a relacionarse con otras variables, es decir, si forman parte de una hipótesis o una teoría. En este caso, se les suele denominar constructos o construcciones hipotéticas.

¿De dónde surgen las hipótesis?

En el enfoque cuantitativo, y si hemos seguido paso por paso el proceso de investigación, es natural que las hipótesis surjan del planteamiento del problema y del marco teórico (de un postulado de una teoría, del análisis de ésta, de generalizaciones empíricas pertinentes a nuestro problema de investigación y de estudios revisados o antecedentes consultados). Existe, pues, una relación muy estrecha entre el planteamiento del problema, la revisión de la literatura y las hipótesis. Al formular las hipótesis volvemos a evaluar nuestro planteamiento del problema.

[2] En esta concepción coinciden diversos autores como Peters (2014), Creswell (2013a), Iversen (2003) y Williams (2003).

Por otra parte, durante el proceso quizá se nos ocurran otras hipótesis que no estaban contempladas en el planteamiento original, producto de nuevas reflexiones, ideas o experiencias; discusiones con profesores, colegas o expertos en el área; incluso, "de analogías, al descubrir semejanzas entre la información referida a otros contextos y la que poseemos para nuestro estudio" (Rojas, 2001). Este último caso ha ocurrido varias veces en las ciencias. Por ejemplo, algunas hipótesis en el área de la comunicación no verbal sobre el manejo de la territorialidad humana surgieron de estudios del tema, pero realizados en animales; algunas concepciones de la teoría del campo o psicología topológica (cuyo principal exponente fue Kurt Lewin) tienen antecedentes en la teoría del comportamiento de los campos electromagnéticos. Las hipótesis de la teoría Galileo —propuestas por Joseph Woelfel y Edward L. Fink (1980)— para medir el proceso de la comunicación, tienen orígenes importantes en la física y otras ciencias exactas (las dinámicas del "yo" se apoyan en nociones del álgebra de vectores).

Las hipótesis pueden surgir por analogía, al aplicar información a otros contextos, como la teoría del campo en psicología, que surgió de la teoría del comportamiento de los campos electromagnéticos. Éstos también forman las auroras boreales.

Asimismo, a veces la experiencia y la observación constante ofrecen materia potencial para el establecimiento de hipótesis importantes, y lo mismo se dice de la intuición. Desde luego, cuanto menor apoyo empírico previo tenga una hipótesis, se deberá tener mayor cuidado en su elaboración y evaluación. No es aceptable formular hipótesis de manera superficial. Establecer hipótesis sin haber revisado cuidadosamente la literatura puede conducirnos a errores como postular algo demasiado comprobado o algo que ha sido rechazado contundentemente. Un ejemplo burdo, pero ilustrativo sería pretender establecer la siguiente hipótesis: "Los seres humanos pueden volar por sí mismos, únicamente con su cuerpo".

¿Qué características debe tener una hipótesis?

1. La hipótesis debe referirse a una situación "real". Como argumenta Castro-Rea (2009), las hipótesis sólo pueden someterse a prueba en un universo y un contexto bien definidos. Por ejemplo, una hipótesis relativa a alguna variable del comportamiento gerencial (digamos, la motivación) deberá someterse a prueba en una situación real (con ciertos gerentes de organizaciones existentes). En ocasiones, en la misma hipótesis se hace explícita esa realidad (por ejemplo, "los niños guatemaltecos que viven en zonas urbanas imitarán más la conducta violenta de la televisión, que los niños guatemaltecos que viven en zonas rurales"), y otras veces la realidad se define por medio de explicaciones que acompañan a la hipótesis. Así, la hipótesis: "cuanto mayor sea la realimentación sobre el desempeño en el trabajo que proporcione un gerente a sus supervisores, más elevada será la motivación intrínseca de éstos hacia sus tareas laborales", no explica qué gerentes, de qué empresas. Y será necesario contextualizar la realidad de dicha hipótesis; afirmar, por ejemplo, que se trata de gerentes de todas las áreas, de empresas exclusivamente industriales con más de mil trabajadores y ubicadas en Medellín, Colombia.

   Es muy frecuente que cuando nuestras hipótesis provienen de una teoría o una generalización empírica (afirmación comprobada varias veces en "la realidad"), sean manifestaciones contextualizadas o casos concretos de hipótesis generales abstractas. La hipótesis: "a mayor satisfacción laboral mayor productividad" es general y susceptible de someterse a prueba en diversas realidades (países, ciudades, parques industriales o aun en una sola empresa; con directivos, secretarias u obreros, etc.; en empresas comerciales, industriales, de servicios o combinaciones de estos tipos, giros o de otras características). En estos casos, al probar nuestra hipótesis contextualizada aportamos evidencia en favor de la hipótesis más general.
2. Las variables o términos de la hipótesis deben ser comprensibles, precisos y lo más concretos que sea posible. Términos vagos o confusos no tienen cabida en una hipótesis. Así, globalización de

la economía y sinergia organizacional son conceptos imprecisos y generales que deben sustituirse por otros más específicos y concretos.

3. La relación entre variables propuesta por una hipótesis debe ser clara y verosímil (lógica). Es indispensable que quede clara la forma en que se relacionan las variables, y esta relación no puede ser ilógica. La hipótesis: "la disminución del consumo del petróleo en Estados Unidos se relaciona con el grado de aprendizaje del álgebra por parte de niños que asisten a escuelas públicas en Buenos Aires", sería inverosímil. No es posible considerarla.
4. Los términos o variables de la hipótesis deben ser observables y medibles, así como la relación planteada entre ellos, o sea, tener referentes en la realidad. Las hipótesis científicas, al igual que los objetivos y las preguntas de investigación, no incluyen aspectos morales ni cuestiones que no podamos medir. Hipótesis como: "los hombres más felices van al cielo" o "la libertad de espíritu está relacionada con la voluntad angelical", implican conceptos o relaciones que no poseen referentes empíricos; por tanto, no son útiles como hipótesis para investigar científicamente ni pueden someterse a prueba en la realidad.
5. Las hipótesis deben estar relacionadas con técnicas disponibles para probarlas. Este requisito está estrechamente ligado con el anterior y se refiere a que al formular una hipótesis, tenemos que analizar si existen técnicas o herramientas de investigación para verificarla, si es posible desarrollarlas y si se encuentran a nuestro alcance.

Se puede dar el caso de que existan esas técnicas, pero que por diversas causas no estén a nuestro alcance. Alguien podría intentar probar hipótesis referentes a la desviación presupuestaria en el gasto gubernamental de un país latinoamericano o a la red de narcotraficantes en la ciudad de Miami, pero no disponer de formas eficaces para obtener sus datos. Entonces, su hipótesis, aunque teóricamente muy valiosa, en la realidad no se puede probar.

¿Qué tipos de hipótesis se pueden establecer?

Hay diversas formas de clasificar las hipótesis, aunque en este apartado nos concentraremos en los siguientes tipos: 2

1. hipótesis de investigación
2. hipótesis nulas
3. hipótesis alternativas
4. hipótesis estadísticas

Veremos las hipótesis estadísticas en el capítulo 8, "Análisis estadístico: segunda parte", del centro de recursos en línea.

Centro de recursos en línea

¿Qué son las hipótesis de investigación?

Lo que a lo largo de este capítulo hemos definido como hipótesis son en realidad las **hipótesis de investigación**, que se definen como proposiciones tentativas acerca de las posibles relaciones entre dos o más variables (Babbie, 2014 y 2012; Martin y Bridgmon, 2012; Davis, 2008; Kalaian y Kasim, 2008 e Iversen, 2003), y que deben cumplir con los cinco requisitos mencionados. Se les suele simbolizar como Hi o H_1, H_2, H_3, etc. (cuando son varias), y también se les denomina "hipótesis de trabajo".

Hipótesis de investigación Proposiciones tentativas sobre las posibles relaciones entre dos o más variables.

Las hipótesis de investigación pueden ser:

a) descriptivas de un valor o dato pronosticado
b) correlacionales
c) de diferencia de grupos
d) causales

Hipótesis descriptivas de un dato o valor que se pronostica[3]

Estas hipótesis se utilizan a veces en estudios descriptivos, para intentar predecir un dato o valor en una o más variables que se van a medir u observar. Pero cabe comentar que no en todas las investigaciones descriptivas se formulan hipótesis de esta clase o que sean afirmaciones más generales ("la ansiedad en los jóvenes alcohólicos será elevada"; "durante este año, los presupuestos de publicidad se incrementarán entre 50 y 70%"; "la motivación extrínseca de los obreros de las plantas de las zonas industriales de Valencia, Venezuela, disminuirá"; "el número de tratamientos psicoterapéuticos aumentará en las urbes sudamericanas con más de tres millones de habitantes"). No es sencillo realizar estimaciones precisas sobre ciertos fenómenos.

Ejemplo

3

Hi: "El aumento del número de divorcios de parejas cuyas edades oscilan entre los 18 y 25 años será de 20% el próximo año" (en un contexto específico como una ciudad o un país).
Hi: "La inflación del próximo semestre no será superior a 3%".

Hipótesis correlacionales

Especifican las relaciones entre dos o más variables y corresponden a los estudios correlacionales ("el tabaquismo está relacionado con la presencia de padecimientos pulmonares"; "la administración de ciertos medicamentos se encuentra asociada con daños físicos a la estructura de los dientes").

Sin embargo, las hipótesis correlacionales no sólo pueden establecer que dos o más variables se encuentran vinculadas, sino también cómo están asociadas. Alcanzan el nivel predictivo y parcialmente explicativo.

En los siguientes ejemplos no sólo se establece que hay relación entre las variables, sino también cómo es la relación (qué dirección sigue). Desde luego, es diferente formular hipótesis en las que dos o más variables están vinculadas, que conjeturar cómo son estas relaciones. En el capítulo 10, "Análisis de los datos cuantitativos", se explica más a fondo el tema de la correlación y los tipos de correlación entre variables.

Ejemplo

"A mayor exposición de los adolescentes a videos musicales con alto contenido sexual, mayor manifestación de estrategias en las relaciones interpersonales para establecer contacto sexual". (Aquí la hipótesis nos indica que cuando una variable aumenta, la otra también y viceversa, cuando una variable disminuye, la otra desciende.)

"A mayor autoestima, habrá menor temor al éxito". (Aquí la hipótesis nos señala que cuando una variable aumenta, la otra disminuye; y si ésta disminuye, aquélla aumenta.)

"Las telenovelas latinoamericanas muestran cada vez mayor contenido sexual en sus escenas". (En esta hipótesis se correlacionan las dos variables siguientes: época o tiempo en que se producen las telenovelas y contenido sexual.)

Es necesario agregar lo siguiente: en una hipótesis de correlación, el orden en que coloquemos las variables no es importante (ninguna variable antecede a la otra; no hay relación de causalidad). Es lo mismo indicar "a mayor X, mayor Y"; que "a mayor Y, mayor X"; o "a mayor X, menor Y"; que "a menor Y, mayor X".

[3] Algunos investigadores consideran a estas hipótesis afirmaciones univariadas. Argumentan que no se relacionan variables. Opinan que, más que relacionar las variables, se está planteando cómo se va a manifestar una variable en una constante (después de todo, el grupo medido de personas u objetos es constante). Este razonamiento tiene cierta validez, por ello, lo dejamos al criterio de cada lector.

Ejemplo 3

"Quienes logran más altas puntuaciones en el examen de estadística tienden a alcanzar las puntuaciones más elevadas en el examen de economía" es igual a: "los que logran tener las puntuaciones más elevadas en el examen de economía son quienes tienden a obtener más altas puntuaciones en el examen de estadística".

Como aprendimos desde pequeños: "el orden de los factores (variables) no altera el producto (la hipótesis)". Desde luego, esto ocurre en la correlación, pero no en las relaciones de causalidad, donde vamos a ver que sí importa el orden de las variables. Pero en la correlación no hablamos de variable independiente (causa) y dependiente (efecto). Cuando sólo hay correlación, estos términos carecen de sentido. Los estudiantes que comienzan en sus cursos de investigación suelen indicar cuál es la variable independiente y cuál la dependiente en toda hipótesis. Ello es un error; únicamente en hipótesis causales se puede hacer esto.

Por otro lado, es común que, cuando en la investigación se pretende correlacionar diversas variables, se tengan varias hipótesis y cada una de ellas relacione un par de variables. Por ejemplo, si quisiéramos relacionar las variables atracción física, confianza, proximidad física y equidad en el noviazgo (todas entre sí), estableceríamos las hipótesis correspondientes.

Ejemplo 3

H_1: "A mayor atracción física, menor confianza"
H_2: "A mayor atracción física, mayor proximidad física"
H_3: "A mayor atracción física, mayor equidad"
H_4: "A mayor confianza, mayor proximidad física"
H_5: "A mayor confianza, mayor equidad"
H_6: "A mayor proximidad física, mayor equidad"

Estas hipótesis deben contextualizarse en su realidad (con qué parejas) y someterse a prueba empírica.

Hipótesis de la diferencia entre grupos

Estas hipótesis se formulan en investigaciones cuya finalidad es comparar grupos. Por ejemplo, supongamos que un publicista piensa que un comercial televisivo en blanco y negro, cuyo objetivo es persuadir a los adolescentes que comienzan a fumar para que dejen de hacerlo, tiene una eficacia diferente que uno a colores. Su pregunta de investigación sería: un comercial de televisión con el mensaje de persuadir a los adolescentes que comienzan a fumar para que dejen de hacerlo, ¿es más eficaz si está en blanco y negro que si está a colores? Y su hipótesis quedaría formulada así:

Ejemplo 3

Hi: "El efecto persuasivo para dejar de fumar no será igual en los adolescentes que vean la versión del comercial televisivo a colores, que el efecto en los adolescentes que vean la versión del comercial en blanco y negro".

Otros ejemplos de este tipo de hipótesis son los siguientes:

Ejemplo 3

Hi: "Los adolescentes le atribuyen más importancia al atractivo físico en sus relaciones de pareja, que las adolescentes a las suyas."
Hi: "El tiempo que tardan en mostrar síntomas de sida las personas contagiadas por transfusión sanguínea, es menor que el de las que adquieren el VIH por transmisión sexual."

En los tres ejemplos anteriores se plantea una posible diferencia entre grupos, sólo que en el primero de ellos únicamente se establece que hay diferencia entre los grupos comparados, pero no se afirma en cuál el impacto será más determinante. No se estipula si el efecto persuasivo es mayor en los adolescentes que ven el comercial en blanco y negro o en quienes lo ven a colores. Se limita a decir que se espera una diferencia. En cambio, en el segundo, la hipótesis, además de establecer la diferencia, especifica cuál de los grupos tendrá mayor valor en la variable de comparación (los jóvenes son quienes, según se piensa, atribuirán mayor importancia al atractivo físico). Lo mismo ocurre en el tercer ejemplo (expresan más lentamente la enfermedad quienes la adquieren por transmisión sexual).

Cuando el investigador no tiene bases para presuponer a favor de qué grupo será la diferencia, formula una hipótesis simple de diferencia de grupos (como en el primer ejemplo de los comerciales). Y cuando sí tiene bases, establece una hipótesis direccional de diferencia de grupos (como en los otros ejemplos). Esto último, por lo común, sucede cuando la hipótesis se deriva de una teoría o estudios antecedentes, o bien, cuando el investigador está bastante familiarizado con el problema de estudio.

Esta clase de hipótesis puede llegar a abarcar dos, tres o más grupos.

Ejemplo

Hi: "Las escenas de la telenovela *La verdad de Paola* presentarán mayor contenido sexual que las de la telenovela *Sentimientos de Christian*, y éstas, a su vez, un mayor contenido sexual que las escenas de la telenovela *Mi último amor, Mariana*".[4]

Algunos investigadores consideran a las hipótesis de diferencia de grupos como un tipo de hipótesis correlacional, porque en última instancia relacionan dos o más variables. El caso del atractivo físico relaciona la variable del género con la variable de la atribución de la importancia del atractivo físico en las relaciones de pareja.

Hipótesis que establecen relaciones de causalidad

Este tipo de hipótesis no solamente afirma la o las relaciones entre dos o más variables y la manera en que se manifiestan, sino que además propone un "sentido de entendimiento" de las relaciones. Tal sentido puede ser más o menos completo, esto depende del número de variables que se incluyan, pero todas estas hipótesis establecen relaciones de causa-efecto.

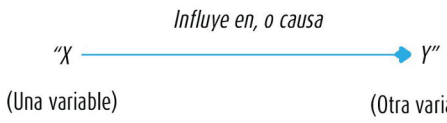
Ejemplo

Hi: "La desintegración del matrimonio provoca baja autoestima en los hijos e hijas". (En el ejemplo, además de establecerse una relación entre las variables, se propone la causalidad de esa relación.)

Hi: "La satisfacción sobre la calidad del diseño ambiental del interior de la oficina donde se labora incrementa significativamente la satisfacción general del espacio de trabajo por parte de sus ocupantes y su desempeño laboral."

Las hipótesis correlacionales pueden simbolizarse como "*X*—*Y*"; y las hipótesis causales, como en la figura 6.1.

Figura 6.1 Simbolización de la hipótesis causal.



[4] Por supuesto, los nombres son ficticios. Si alguna telenovela se ha titulado o se titulará así, es una coincidencia.

Correlación y causalidad son conceptos asociados, pero distintos. Si dos variables están correlacionadas, ello no necesariamente implica que una será causa de la otra. Supongamos que una empresa fabrica un producto que se vende poco y decide mejorarlo. Entonces, lanza una campaña para anunciar el producto en radio y televisión. Después, se observa un aumento en las ventas del producto. Los ejecutivos de la empresa pueden decir que el lanzamiento de la campaña está relacionado con el incremento de las ventas; pero si no se demuestra la causalidad, no es posible asegurar que la campaña haya provocado tal incremento. Quizá la campaña sea la causa del aumento, pero tal vez la causa sea en sí la mejora al producto, una excelente estrategia de comercialización u otro factor, o bien todas pueden ser las causas.

Otro caso es el que se explicó en el capítulo anterior, cuando la estatura parecía estar correlacionada con la inteligencia en los niños (los de mayor estatura sacaban mejores calificaciones en la prueba de inteligencia). Pero la realidad fue que la maduración era la variable que estaba relacionada con la respuesta a una prueba de inteligencia (más que a la inteligencia en sí). La correlación no tenía sentido; mucho menos lo tendría establecer una causalidad, al afirmar que la estatura es causa de la inteligencia o que, por lo menos, influye en ella. Es decir, no todas las correlaciones tienen sentido y no siempre que se encuentra una correlación puede inferirse causalidad. Si cada vez que se obtiene una correlación se asumiera la causalidad, ello equivaldría a decir que cada vez que se observa a una señora y a un niño juntos se supusiera que ella es su madre, cuando puede ser su tía, una vecina o una señora que por azar se colocó muy cerca del chico.

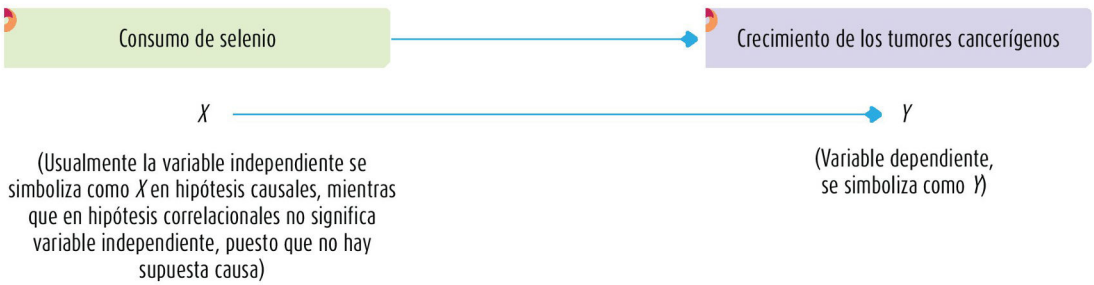
Para establecer causalidad, primero debe haberse demostrado correlación, pero además la causa debe ocurrir antes que el efecto. Asimismo, los cambios en la causa tienen que provocar cambios en el efecto.

Al hablar de hipótesis, a las *supuestas causas* se les conoce como *variables independientes* y a los *efectos* como *variables dependientes*. Únicamente es posible hablar de variables independientes y dependientes cuando se formulan hipótesis causales o hipótesis de la diferencia de grupos, siempre y cuando en estas últimas se explique cuál es la causa de la diferencia supuesta en la hipótesis.

A continuación se exponen distintos tipos de hipótesis causales:

1. **Hipótesis causales bivariadas**. En éstas se plantea una relación entre una variable independiente y una variable dependiente. Por ejemplo: "el consumo diario y permanente de selenio como suplemento alimenticio reduce el crecimiento de los tumores cancerígenos en mujeres que se encuentran en la etapa inicial de la enfermedad" (vea la figura 6.2).

Figura 6.2 Esquema de relación causal bivariada.

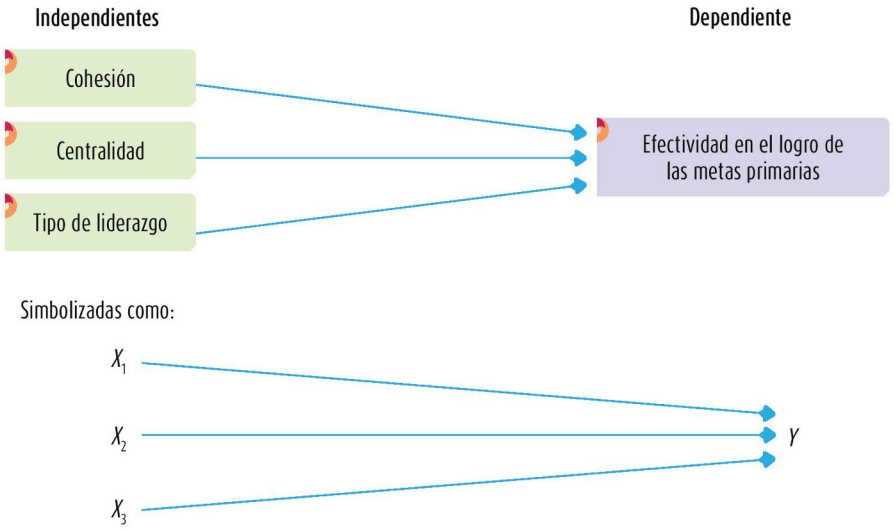


2. **Hipótesis causales multivariadas**. Plantean una relación entre diversas variables independientes y una dependiente, una independiente y varias dependientes o diversas variables independientes y varias dependientes.

Ejemplo

"La cohesión y la centralidad en un grupo sometido a una dinámica, así como el tipo de liderazgo que se ejerza dentro del grupo, determinan la eficacia de éste para alcanzar sus principales metas" (figura 6.3).

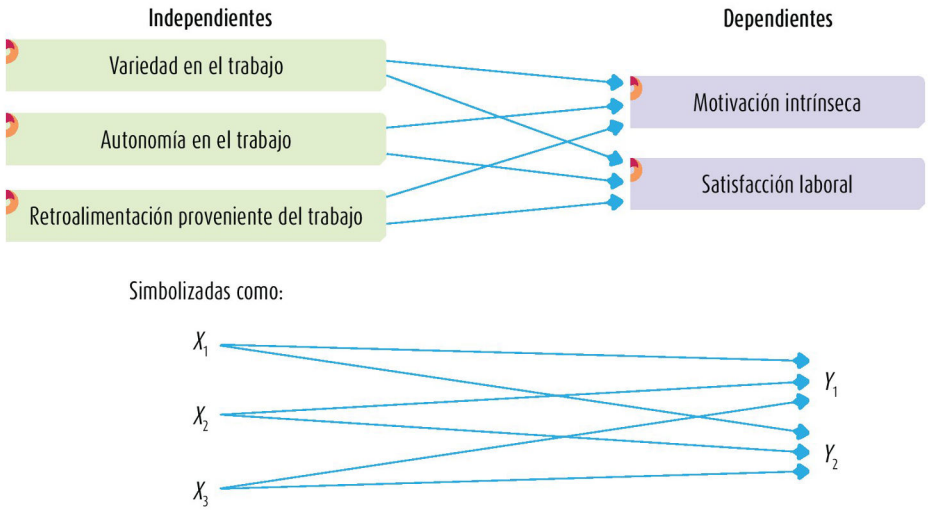
Figura 6.3 Esquema de relación causal multivariada.



Ejemplo

"La variedad y la autonomía en el trabajo, así como la realimentación proveniente del desarrollo de éste, generan mayor motivación intrínseca y satisfacción laborales" (figura 6.4).

Figura 6.4 Esquema de relación causal multivariada con dos variables dependientes.



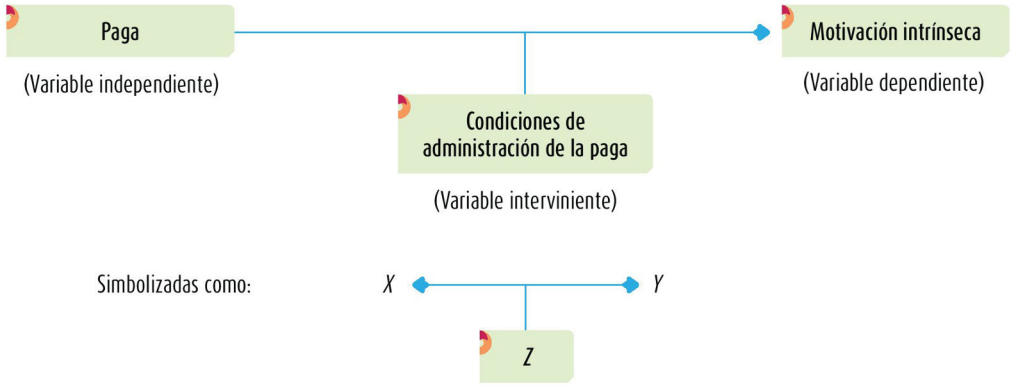
Otro ejemplo sería: "El momento de flexión residual y la resistencia a la cizalladura en las vigas de concreto reforzado dañadas por los incendios son factores que determinan la seguridad de la estructura" (dos variables independientes y una dependiente).

Las hipótesis multivariadas pueden plantear otro tipo de relaciones causales, en las que ciertas variables intervienen y modifican la relación (hipótesis con presencia de variables intervinientes).

Ejemplo

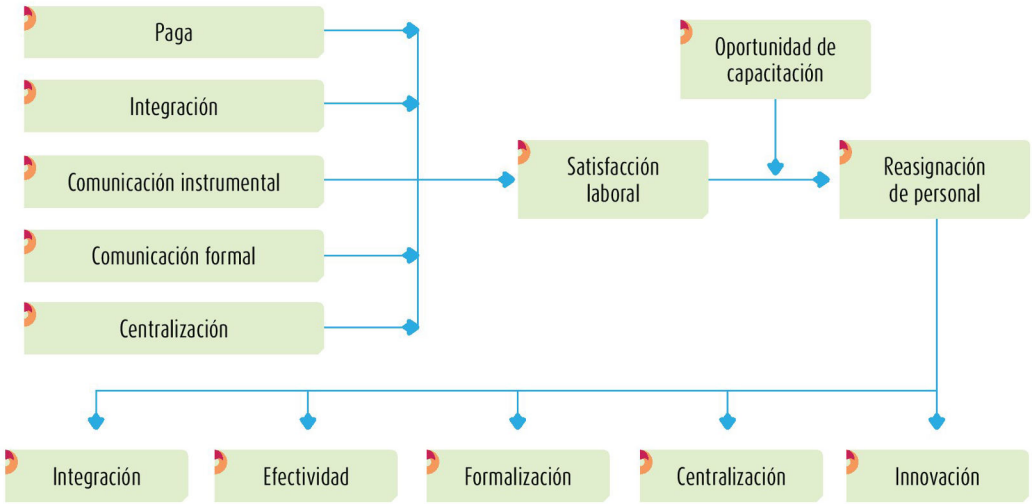
"La paga aumenta la motivación intrínseca de los trabajadores, cuando se administra de acuerdo con el desempeño" (figura 6.5).

Figura 6.5 Esquema causal con variable interviniente.



Es posible que haya estructuras causales de variables más complejas que resulta difícil expresar en una sola hipótesis, porque las variables se relacionan entre sí de distintas maneras. Entonces se plantean las relaciones causales en dos o más hipótesis, o de forma gráfica (vea figura 6.6).

Figura 6.6 Estructura causal compleja multivariada.



La figura 6.6[5] podría desglosarse en múltiples hipótesis; por ejemplo:

H_1: "La paga incrementa la satisfacción laboral".
H_2: "La integración, la comunicación instrumental y la comunicación formal incrementan la satisfacción laboral".
H_3: "La centralización disminuye la satisfacción laboral".
H_4: "La satisfacción laboral influye en la reasignación de personal".
H_5: "La oportunidad de capacitación mediatiza la vinculación entre la satisfacción laboral y la reasignación de personal".
H_6: "La reasignación de personal afecta la integración, la efectividad organizacional, la formalización, la centralización y la innovación".

Cuando las hipótesis causales se someten al análisis estadístico, se evalúa la influencia de cada variable independiente (causa) en la dependiente (efecto), y la influencia conjunta de todas las variables independientes en la dependiente o dependientes.

[5] Las variables fueron extraídas de Price (1977) y Hernández-Sampieri (2005).

¿Qué son las hipótesis nulas?[6]

Hipótesis nulas Proposiciones que niegan o refutan la relación entre variables.

Las **hipótesis nulas** son, en cierto modo, el reverso de las hipótesis de investigación. También constituyen proposiciones acerca de la relación entre variables, sólo que sirven para refutar o negar lo que afirma la hipótesis de investigación.[7] Si la hipótesis de investigación propone: "los adolescentes le atribuyen más importancia al atractivo físico en sus relaciones de pareja que las adolescentes", la hipótesis nula postularía: "los adolescentes *no* le atribuyen más importancia al atractivo físico en sus relaciones de pareja que las adolescentes".

Debido a que este tipo de hipótesis resulta la contrapartida de la hipótesis de investigación, hay prácticamente tantas clases de hipótesis nulas como de investigación. Es decir, la clasificación de hipótesis nulas es similar a la tipología de las hipótesis de investigación: hipótesis nulas descriptivas de un valor o dato pronosticado, hipótesis que niegan o contradicen la relación entre dos o más variables, hipótesis que niegan que haya diferencia entre grupos que se comparan e hipótesis que niegan la relación de causalidad entre dos o más variables (en todas sus formas). Las hipótesis nulas se simbolizan así: Ho.

Veamos algunos ejemplos de hipótesis nulas que corresponden a ejemplos de hipótesis de investigación que se mencionaron.

Ejemplo

Ho: "El aumento del número de divorcios de parejas cuyas edades oscilan entre los 18 y 25 años, no será de 20% el próximo año."

Ho: "La administración de medicamentos no se encuentra asociada con daños físicos a la estructura de los dientes."

Ho: "Las escenas de la telenovela *La verdad de Paola* no presentarán mayor contenido sexual que las de la telenovela *Sentimientos de Christian*, ni éstas tendrán mayor contenido sexual que las escenas de la telenovela *Mi último amor, Mariana.*" Esta hipótesis niega la diferencia entre grupos y también podría formularse así: "No existen diferencias en el contenido sexual entre las escenas de las telenovelas *La verdad de Paola*, *Sentimientos de Christian* y *Mi último amor, Mariana*". O bien: "El contenido sexual de las telenovelas *La verdad de Paola*, *Sentimientos de Christian* y *Mi último amor, Mariana* es el mismo".

Ho: "La satisfacción sobre la calidad del diseño ambiental del interior de la oficina donde se labora no incrementa la satisfacción general del espacio de trabajo por parte de sus ocupantes ni su desempeño laboral" (hipótesis que niega la relación causal).

¿Qué son las hipótesis alternativas?

Como su nombre lo indica, son posibilidades alternas de las hipótesis de investigación y nula: ofrecen una descripción o explicación distinta de las que proporcionan éstas. Si la hipótesis de investigación establece: "esta silla es roja", la nula afirmará: "esta silla no es roja", y podrían formularse una o más hipótesis alternativas: "esta silla es azul", "esta silla es verde", "esta silla es amarilla", etc. Cada una constituye una descripción distinta de las que proporcionan las hipótesis de investigación y nula.

Hipótesis alternativas Son posibilidades diferentes o "alternas" de las hipótesis de investigación y nula.

Las **hipótesis alternativas** se simbolizan como Ha y sólo pueden formularse cuando efectivamente hay otras posibilidades, además de las hipótesis de investigación y nula. De no ser así, no deben establecerse.

[6] El sentido que en este libro se da a la hipótesis nula es el más común, el de negación de la hipótesis de investigación (Babbie, 2012; Sullivan, 2009; Lavrakas, 2008 y Voi, 2003), el cual fue propuesto por Fisher (1925). No se plantean otras connotaciones o usos del término (por ejemplo, especificar un parámetro de cero) porque se generarían confusiones entre estudiantes que se inician en la investigación. Para aquellos que deseen conocer más del tema, se recomiendan las siguientes fuentes: Krueger (2006), Van Dalen y Meyer (1994, pp. 403-404) y, sobre todo, Henkel (1976, pp. 34-40).

[7] La hipótesis nula es un componente esencial de la prueba de hipótesis en la investigación. Es relevante cuando se efectúan mediciones y las hipótesis han sido derivadas de teorías y tienen que ser probadas. La hipótesis de investigación define cierto patrón que se encontrará en los datos, y el análisis estadístico se diseña para evaluar el grado al cual la evidencia de las medidas recogidas apoya la existencia de ese patrón. La hipótesis nula es la hipótesis que indica que el patrón encontrado en los datos simplemente se debe a la casualidad (Krueger, 2006; Voi, 2003).

Ejemplo

Hi: "El candidato A obtendrá en la elección para la presidencia del consejo escolar entre 50 y 60% de la votación total".
Ho: "El candidato A no obtendrá en la elección para la presidencia del consejo escolar entre 50 y 60% de la votación total".
Ha: "El candidato A obtendrá en la elección para la presidencia del consejo escolar más de 60% de la votación total".
Ha: "El candidato A obtendrá en la elección para la presidencia del consejo escolar menos de 50% de la votación total".

Ejemplo

Hi: "Los jóvenes le atribuyen más importancia al atractivo físico en sus relaciones de pareja que las jóvenes".
Ho: "Los jóvenes no le atribuyen más importancia al atractivo físico en sus relaciones de pareja que las jóvenes".
Ha: "Los jóvenes le atribuyen menos importancia al atractivo físico en sus relaciones de pareja que las jóvenes".

En este último ejemplo de los jóvenes, si la hipótesis nula hubiera sido formulada de la siguiente manera:

Ejemplo

Ho: "Los jóvenes no le atribuyen más importancia o le atribuyen menos importancia al atractivo físico en sus relaciones de pareja que las jóvenes".

No habría posibilidad de formular una hipótesis alternativa, puesto que las hipótesis de investigación y nula abarcan todas las posibilidades.

Las hipótesis alternativas, como puede verse, constituyen otras hipótesis de investigación además de la hipótesis de investigación original.

¿En una investigación se formulan hipótesis de investigación, nula y alternativa?

Al respecto no hay reglas universales, ni siquiera consenso entre los investigadores. Se puede leer un artículo de alguna revista científica en la que sólo se formule la hipótesis de investigación, y podemos encontrar un artículo en otra revista en la que únicamente se plantea la hipótesis nula. Asimismo, podemos descubrir artículos en una tercera revista, en los cuales se establezcan las hipótesis de investigación y nula, pero no las alternativas y, en fin, en una cuarta publicación podemos ver otro artículo que contenga las hipótesis de investigación, nulas y alternativas, etc. Esta situación es similar en reportes académicos (como tesis) y no académicos.

La opción tal vez más común es incluir únicamente la o las hipótesis de investigación (Degelman, 2005, consultor de la American Psychological Association). Algunos investigadores sólo enuncian la hipótesis nula o de investigación presuponiendo que quien lea su reporte deducirá la hipótesis contraria.

La American Psychological Association (APA, 2011) recomienda, para decidir qué tipo de hipótesis deben incluirse en el informe, que se consulten los manuales o a un asesor calificado de su universidad o las normas de publicaciones de los estilos APA, Vancouver, Harvard y otros, los cuales se comentan en el capítulo 11 de esta obra.

Desde luego, siempre se tienen presentes todos los tipos de hipótesis aunque sólo se escriba uno.

¿Cuántas hipótesis se deben formular en una investigación?

Cada investigación es diferente. Algunas contienen gran variedad de hipótesis porque el problema de investigación es complejo (por ejemplo, pretenden relacionar 15 o más variables), mientras que otras contienen una o dos hipótesis. Todo depende del planteamiento del problema.

La calidad de una investigación no está relacionada con el número de hipótesis que contenga. En este sentido, se debe tener el número de hipótesis necesarias para guiar el estudio, ni una más ni una menos.

¿En una investigación se pueden formular hipótesis descriptivas de un dato que se pronostica en una variable, hipótesis correlacionales, hipótesis de la diferencia de grupos e hipótesis causales?

La respuesta es *sí*. En una misma investigación es posible establecer todos los tipos de hipótesis, porque el problema de investigación así lo requiere. Supongamos que alguien ha planteado un estudio en una ciudad latinoamericana y sus preguntas de investigación e hipótesis son las que se muestran en la tabla 6.2.

Tabla 6.2 Ejemplo de un estudio con varias preguntas de investigación e hipótesis

Preguntas de investigación	Hipótesis
¿Cuál será a fin de año el nivel de desempleo en la ciudad de Baratillo?	El nivel de desempleo en la ciudad de Baratillo será de 5% a fin de año (Hi: % = 5).
¿Cuál es el nivel promedio de ingreso familiar mensual en la ciudad de Baratillo?	El nivel promedio de ingreso familiar mensual oscila entre 650 y 700 dólares (Hi: $649 < \bar{X} < 701$).
¿Existen diferencias entre los distritos (barrios, delegaciones o equivalentes) de la ciudad de Baratillo en cuanto al nivel de desempleo? (¿Hay barrios o distritos con mayores índices de desempleo?)	Existen diferencias en cuanto al nivel de desempleo entre los distritos de la ciudad de Baratillo (Hi: Índice 1 ≠ Índice 2 ≠ Índice 3 ≠ Índice *k*).
¿Cuál es el nivel de escolaridad promedio de los jóvenes y las jóvenes que viven en Baratillo? ¿Existen diferencias por género al respecto?	No se dispone de información, no se establecen hipótesis.
¿Está relacionado el desempleo con incrementos en la delincuencia de dicha ciudad?	A mayor desempleo, mayor delincuencia (Hi: $r_{xy} \neq 0$).
¿Provoca el nivel de desempleo un rechazo de la política fiscal gubernamental?	El desempleo provoca un rechazo contra la política fiscal gubernamental (Hi: $X \rightarrow Y$).

En el ejemplo encontramos todos los tipos generales de hipótesis. Asimismo, observaremos que hay preguntas que no se traducen en hipótesis (escolaridad y diferencias por género en ésta). Ello puede deberse a que es difícil establecerlas, ya que no se dispone de información al respecto.

Los estudios que se inician y concluyen como descriptivos, formularán —si pronostican un dato— hipótesis descriptivas; los correlacionales podrán establecer hipótesis descriptivas de estimación, correlacionales y de diferencia de grupos (cuando éstas no expliquen la causa que provoca la diferencia); por su parte, los explicativos podrán incluir hipótesis descriptivas de pronóstico, correlacionales, de diferencia de grupos y causales. No debemos olvidar que una investigación puede abordar parte del problema de forma descriptiva y parte explicativa. Aunque debemos señalar que los estudios descriptivos no suelen contener hipótesis, y ello se debe a que en ocasiones es difícil precisar el valor que se puede manifestar en una variable.

Los tipos de estudio que no establecen hipótesis son los exploratorios. No puede presuponerse (afirmando) algo que apenas va a explorarse. Sería como si antes de una primera cita con una persona

totalmente desconocida del género opuesto, trataríamos de conjeturar qué tan simpática es, qué intereses y valores tiene, etc. Ni siquiera podríamos anticipar qué tan atractiva nos va a resultar, y tal vez en una primera cita nos dejemos llevar por nuestra imaginación. Pero en una investigación esto no debe ocurrir. Si se nos proporciona más información (lugares a donde le agrada ir, ocupación, religión, nivel socioeconómico, tipo de música que le gusta y grupos de los que es miembro), podemos plantearnos hipótesis en mayor medida, aunque nos basemos en estereotipos. Y si nos dieran información muy personal e íntima sobre ella, podríamos sugerir hipótesis acerca de qué clase de relación vamos a establecer con esa persona y por qué (explicaciones tentativas).

¿Qué es la prueba de hipótesis?

Como se ha dicho, en el proceso cuantitativo las hipótesis se someten a prueba o escrutinio empírico para determinar si son apoyadas o refutadas, de acuerdo con lo que el investigador observa. De hecho, para esto se formulan en la tradición deductiva. Ahora bien, en realidad no podemos probar que una hipótesis sea verdadera o falsa, sino argumentar que fue apoyada o no de acuerdo con ciertos datos obtenidos en una investigación particular. Desde el punto de vista técnico, no se acepta una hipótesis por medio de un estudio, sino que se aporta evidencia a favor o en contra.[8] Cuantas más investigaciones apoyen una hipótesis, más credibilidad tendrá y, por supuesto, será válida para el contexto (lugar, tiempo y participantes, casos o fenómenos) en que se comprobó. Al menos lo es probabilísticamente.

Las hipótesis, en el enfoque cuantitativo, se someten a prueba en la "realidad" cuando se implementa un diseño de investigación, se recolectan datos con uno o varios instrumentos de medición, y se analizan e interpretan esos mismos datos.

¿Cuál es la utilidad de las hipótesis?

Es posible que alguien piense que con lo expuesto en este capítulo queda claro qué valor tienen las hipótesis para la investigación. Sin embargo, creemos que es necesario ahondar un poco más en este punto, mencionando las principales funciones de las hipótesis.

1. En primer lugar, son las guías de una investigación en el enfoque cuantitativo. Formularlas nos ayuda a saber lo que tratamos de buscar, de probar. Proporcionan orden y lógica al estudio. Son como los objetivos de un plan administrativo: las sugerencias formuladas en las hipótesis pueden ser soluciones a los problemas de investigación. Si de hecho lo son o no, es la tarea del estudio (Castro-Rea, 2009).
2. En segundo lugar, tienen una función descriptiva y explicativa, según sea el caso. Cada vez que una hipótesis recibe evidencia empírica a favor o en contra, nos dice algo acerca del fenómeno con el que se asocia o hace referencia. Si la evidencia es a favor, la información sobre el fenómeno se incrementa; y aun si la evidencia es en contra, descubrimos algo acerca del fenómeno que no sabíamos antes.
3. La tercera función es probar teorías. Cuando varias hipótesis de una teoría reciben evidencia positiva, la teoría va haciéndose más sólida; y cuanto más evidencia haya a favor de aquéllas, más evidencia habrá a favor de ésta.
4. Una cuarta función consiste en sugerir teorías. Diversas hipótesis no están asociadas con ninguna teoría; pero llega a suceder que como resultado de la prueba de una hipótesis, se pueda formular una teoría o las bases para ésta.

[8] Aquí se ha preferido evitar la exposición sobre la lógica de la prueba de hipótesis, la cual indica que la única alternativa abierta en una prueba de significancia para una hipótesis, radica en que se puede rechazar una hipótesis nula o equivocarse al rechazarla. Pero la frase "equivocarse al rechazar" no es sinónimo de aceptar. La razón para no incluir esta perspectiva reside en que, el hacerlo, podría confundir más que esclarecer el panorama al que se inicia en el tema. A quien desee ahondar en la lógica de la prueba de hipótesis, le recomendamos acudir a Martin y Bridgmon (2012), Blaikie (2007 y 2000), Chalmers (1999) y especialmente a Henkel (1976, pp. 34-35), así como otras referencias que sustentan desde la epistemología las posiciones al respecto, como Popper (1992 y 1996) y Hanson (1958).

¿Qué ocurre cuando no se aporta evidencia a favor de las hipótesis de investigación?

No es raro escuchar una conversación como la siguiente entre dos pasantes que acaban de analizar los datos de su tesis (que es una investigación):

Elisa: Los datos no apoyan nuestras hipótesis.
Gabriel: ¿Y ahora qué vamos a hacer? Nuestra tesis no sirve.
Elisa: Tendremos que hacer otra tesis.

No siempre los datos y resultados apoyan las hipótesis. Pero el hecho de que éstos no aporten evidencia en favor de las hipótesis planteadas de ningún modo significa que la investigación carezca de utilidad. Recordemos que en la investigación el fin último es el conocimiento y, en este sentido, también los datos en contra de una hipótesis ofrecen entendimiento. Lo importante es analizar por qué no se aportó evidencia a favor de las hipótesis.

A propósito, conviene citar a Van Dalen y Meyer (1994, p. 193):

> Para que las hipótesis tengan utilidad, no es necesario que sean las respuestas correctas a los problemas planteados. En casi todas las investigaciones, el estudioso formula varias hipótesis y espera que alguna de ellas proporcione una solución satisfactoria del problema. Al eliminar cada una de las hipótesis, va estrechando el campo en el cual deberá hallar la respuesta.

Y agregan:

> La prueba de "hipótesis falsas" [que nosotros preferimos llamar *hipótesis que no recibieron evidencia empírica*] también resulta útil si dirige la atención del investigador hacia factores o relaciones insospechadas que, de alguna manera, podrían ayudar a resolver el problema.

La American Psychological Association (2011) señala que en los reportes o informes de resultados se mencionen todos los resultados relevantes, incluyendo aquellos que contradigan las hipótesis.

¿Deben definirse las variables de una hipótesis como parte de su formulación?

Al formular una hipótesis, es indispensable definir **los términos o variables** incluidos en ella. Esto es necesario por varios motivos:

1. Para que el investigador, sus colegas, los lectores del estudio y, en general, cualquier persona que consulte la investigación le dé el mismo significado a los términos o variables de las hipótesis, es común que un mismo concepto se emplee de maneras distintas. El término "novios" puede significar para alguien una relación entre dos personas de género opuesto que se comunican con la mayor frecuencia que les es posible, que cuando están juntos se besan y se toman de la mano, que se sienten atraídos en lo físico y comparten entre sí información que nadie más posee. Para otros significaría una relación entre dos personas de género diferente que tiene como finalidad contraer matrimonio. Para un tercero, una relación entre dos individuos de género distinto que mantienen relaciones íntimas, y alguien más podría tener otra concepción. Y en caso de que se pensara llevar a cabo un estudio con parejas de novios, no sabríamos con exactitud quiénes se incluirían en éste y quiénes no, a menos que se definiera con la mayor precisión posible el concepto de "novios". Términos como "actitud", "inteligencia" y "aprovechamiento" llegan a tener varios significados o definirse de diversas formas.
2. Asegurarnos de que las variables pueden ser medidas, observadas, evaluadas o inferidas, es decir, que de ellas se pueden obtener datos en la realidad.
3. Confrontar nuestra investigación con otras similares. Si tenemos definidas nuestras variables, podemos comparar nuestras definiciones con las de otros estudios para saber "si hablamos de lo mismo". Si la comparación es positiva, confrontaremos los resultados de nuestra investigación con los resultados de las demás.

www.elosopanda.com | jamespoetrodriguez.com

4. Evaluar más adecuadamente los resultados de nuestra investigación, porque las variables, y no sólo las hipótesis, se contextualizan.

En conclusión, sin definición de las variables no hay investigación. Las variables deben ser definidas de dos formas: conceptual y operacionalmente.

Definición conceptual o constitutiva

Una **definición conceptual** trata a la variable con otros términos. Así, *trastorno bipolar* se podría definir como una enfermedad mental caracterizada por cambios drásticos de humor o estado de ánimo, de muy alto (fase maníaca) a muy bajo (fase depresiva) (Bryant, 2007); y *poder* como: "influir más en los demás de lo que éstos influyen en uno". Se tratan de definiciones de diccionarios o de libros especializados (Kerlinger y Lee, 2002), y cuando describen la esencia o las características de una variable, objeto o fenómeno se les denomina definiciones reales (Reynolds, 1986). Estas últimas constituyen la adecuación de la definición conceptual a las necesidades prácticas de la investigación. De esa forma, el término actitud se definiría como "una tendencia o predisposición a evaluar de cierta manera un objeto o un símbolo de este objeto" (Perloff, 2013; Maio y Haddock, 2010; y Oskamp y Schultz, 2005). Si nuestra hipótesis fuera: "Cuanto mayor sea la exposición de los votantes indecisos a entrevistas televisivas concedidas por los candidatos contendientes, más favorable será la actitud hacia el acto de votar", tendríamos que contextualizar la definición conceptual de "actitud" (formular la definición real). La "actitud hacia el acto de votar" podría definirse como la predisposición a evaluar como positivo el hecho de votar para una elección.

Algunos ejemplos de definiciones conceptuales se muestran en la tabla 6.3

Tabla 6.3 Ejemplos de definiciones conceptuales

Variable	Definición conceptual
Inteligencia emocional	Capacidad para reconocer y controlar nuestras emociones, así como manejar con más destreza nuestras relaciones (Goleman, 1996).
Producto interno bruto	Conjunto del valor de todos los bienes y servicios finales producidos en una economía durante un periodo determinado, que puede ser trimestral o anual. El PIB puede ser clasificado como nominal o real. En el primero, los bienes y servicios finales son valuados a los precios vigentes durante el periodo en cuestión, mientras que en el segundo los bienes y servicios finales se valúan a los precios vigentes en un año base (Centro de Investigación y Docencia Económicas, 2004).
Abuso sexual infantil	La utilización de un menor de 12 años o menos para la satisfacción sexual de un adulto. El abuso sexual en la niñez puede incluir contacto físico, masturbación, relaciones sexuales (incluso penetración) y/o contacto anal u oral. Pero también puede incluir el exhibicionismo, voyeurismo, la pornografía y/o la prostitución infantil (International Planned Parenthood Federation, 2000).
Clima organizacional	Conjunto de percepciones compartidas por los empleados respecto de factores de su entorno laboral (Hernández-Sampieri, Méndez y Contreras, 2013).
Pareja ideal (en las relaciones románticas)	Prototipo de ser humano que los individuos consideran que posee los atributos más valorados por ellos y que representaría la opción perfecta para implicarse en una relación amorosa romántica e íntima de largo plazo (casarse o al menos vivir con ella) (Hernández-Sampieri y Mendoza, 2008).

Tales definiciones son necesarias pero insuficientes para definir las variables de la investigación, porque no nos vinculan directamente con "la realidad" o con "el fenómeno, contexto, expresión, comunidad o situación". Después de todo, continúan con su carácter de conceptos. Los científicos necesitan ir más allá, deben definir las variables que se utilizan en sus hipótesis, en forma tal que puedan ser comprobadas y contextualizadas. Lo anterior es posible por medio de lo que se conoce como definiciones operacionales.

Definiciones operacionales

Definición operacional Conjunto de procedimientos y actividades que se desarrollan para medir una variable.

Una **definición operacional** constituye el conjunto de procedimientos que describe las actividades que un observador debe realizar para recibir las impresiones sensoriales, las cuales indican la existencia de un concepto teórico en mayor o menor grado (Reynolds, 1986, p. 52). En otras palabras, especifica qué actividades u operaciones deben realizarse para medir una variable e interpretar los datos obtenidos (Hernández-Sampieri *et al.*, 2013). Una definición operacional nos dice que para recoger datos respecto de una variable, hay que hacer esto y esto otro, además articula los procesos o acciones de un concepto que son necesarios para identificar ejemplos de éste (MacGregor, 2006). Así, la definición operacional de la variable "temperatura" sería el termómetro y la escala elegida (por ejemplo, grados centígrados); e "inteligencia" se definiría operacionalmente como las respuestas a una determinada prueba de inteligencia (por ejemplo: Stanford-Binet o Wechsler). Con respecto a la "satisfacción sexual de adultos", hay varias definiciones operacionales para medir este constructo: la escala de satisfacción con la vida sexual (*The Satisfaction With Sex Life Scale,* SWSLS) (Neto, 2012); Inventario de satisfacción sexual de Golombok y Rust (*Golombok Rust Inventory of Sexual Satisfaction*, GRISS) (Rust y Golombok, 1986; Meston y Derogatis, 2002) y el inventario de satisfacción sexual (Álvarez-Gayou Jurgenson *et al.*, 2005),[9] para ambos géneros.

La variable "ingreso familiar" podría operacionalizarse al preguntar sobre el ingreso personal de cada uno de los miembros de la familia y luego sumar las cantidades que cada quien indicó. El atractivo físico en un certamen de belleza se operacionaliza al aplicar una serie de criterios que un jurado utiliza para evaluar a las candidatas; los miembros del jurado otorgan una calificación a las contendientes en cada criterio y después obtienen una puntuación total del atractivo físico.

Casi siempre se dispone de varias definiciones operacionales (o formas de operacionalizar) de una variable. Para definir operacionalmente la variable "personalidad" se cuenta con diversas alternativas: las pruebas psicométricas, como las diferentes versiones del Inventario Multifacético de la Personalidad Minnesota (MMPI); pruebas proyectivas como el test de Rorschach o el test de apercepción temática (TAT), etcétera.

Es posible medir la ansiedad de un individuo por medio de la observación directa de los expertos, quienes juzgan el nivel de ansiedad de esa persona; con mediciones fisiológicas de la actividad del sistema psicológico (presión sanguínea, respiraciones, etc.) y con el análisis de las respuestas a un cuestionario de ansiedad (Reynolds, 1986, p. 52). El aprendizaje de un alumno en un curso de investigación se mediría con el empleo de varios exámenes, un trabajo, o una combinación de exámenes, trabajos y prácticas.

Algunos ejemplos de definiciones operacionales se incluyen en la tabla 6.4 (se muestran únicamente los nombres y algunas características).

Tabla 6.4 Ejemplos de definiciones operacionales

Variable	Definición operacional
Inteligencia emocional	EIT (*Emotional Intelligence Test*). Prueba con 70 ítems o reactivos.
Aceleración	Acelerómetro con su escala en unidades "g".
Abuso sexual infantil	*Children's Knowledge of Abuse Questionnaire-Revised* (CKAQ-R). Versión en español. El CKAQ-R tiene 35 preguntas a responder como verdadero-falso, y cinco extras para ser administradas a niñas y niños de ocho años en adelante. Puede ser aplicado a cualquier infante sin previa instrucción.
Clima organizacional	Escala Clima-UNI con 73 ítems para medir las siguientes dimensiones del clima organizacional: moral, apoyo de la dirección, innovación, percepción de la empresa-identidad-identificación, comunicación, percepción del desempeño, motivación intrínseca, autonomía, satisfacción general, liderazgo, visión y recompensas o retribución.

(*continúa*)

Centro de recursos en línea

[9] El desarrollo de esta definición operacional de satisfacción sexual lo podrá encontrar el lector en el ejemplo 4 del centro de recursos en línea y descargarlo (en Material complementario → Ejemplos → Diseño de una escala autoaplicable para la evaluación de la satisfacción sexual en hombres y mujeres mexicanos).

Tabla 6.4 (*continuación*)

Variable	Definición operacional
Satisfacción con la calidad de los muebles de oficina	Grado de satisfacción con: *a*) el confort del mobiliario personal para trabajar (silla, escritorio, computadora, equipo, etc.); *b*) la capacidad de ajuste y adaptabilidad de los muebles y *c*) los colores y texturas del piso, los muebles y los acabados de las superficies (mediante escalas de tipo Likert: "muy satisfecho", "satisfecho", "insatisfecho" y "muy insatisfecho" en cada rubro, las cuales se explican en el capítulo 9 de este libro: "Recolección de los datos cuantitativos") (Lee y Guerin, 2009).
Crecimiento de los tumores cancerígenos	Mamografías comparativas y, recientemente, con modelos de una función continua del tamaño del tumor que consideran volumen, diámetro y tiempo, así como edad y datos de la población donde se efectúa el estudio (ritmo de crecimiento) (Weedon-Fekjær, Lindqvist, Vatten, Aalen y Tretli, 2008).

Cuando el investigador dispone de varias opciones para definir operacionalmente una variable, debe elegir la que proporcione mayor información sobre la variable, capte mejor su esencia, se adecue más a su contexto y sea más precisa. O bien, una mezcla de tales alternativas.

Los criterios para evaluar una definición operacional son básicamente cuatro: adecuación al contexto, capacidad para captar los componentes de la variable de interés, confiabilidad y validez. De ellos se hablará en el capítulo 9, "Recolección de los datos cuantitativos". Una correcta selección de las definiciones operacionales disponibles o la creación de la propia definición operacional se encuentran muy relacionadas con una adecuada revisión de la literatura. Cuando ésta ha sido cuidadosa, se tiene una gama más amplia de definiciones operacionales para elegir o más ideas para desarrollar una nueva. Asimismo, al contar con estas definiciones, el tránsito a la elección del o los instrumentos para recabar los datos es rápido, pues sólo debemos considerar que se adapten al diseño y a la muestra del estudio.

En una investigación se tienen por lo regular diversas variables y, por tanto, se formularán varias definiciones conceptuales y operacionales.

Algunas variables no requieren que su definición conceptual se mencione en el reporte de investigación, porque ésta es relativamente obvia y compartida. El mismo título de la variable la define; por ejemplo, "género" y "edad". Pero prácticamente todas las variables requieren una definición operacional para ser evaluadas de manera empírica, aun cuando en el estudio no se formulen hipótesis. Siempre que se tengan variables, se deben definir operacionalmente. En el siguiente ejemplo se muestra una hipótesis con las correspondientes definiciones operacionales de las variables que la integran.

Ejemplo

Hi: "A mayor motivación intrínseca en el trabajo, menor ausentismo."

Variable =	"Motivación intrínseca en el trabajo".	"Ausentismo laboral".
Definiciones conceptuales:	"Estado cognitivo que refleja el grado en que un trabajador atribuye la fuerza de su comportamiento en el trabajo a satisfacciones o beneficios derivados de sus tareas laborales en sí mismas. Es decir, a sucesos que no están mediatizados por una fuente externa a las tareas laborales del trabajador. Este estado de motivación puede ser señalado como una experiencia autosatisfactoria".	"El grado en el cual un trabajador no se reporta a trabajar a la hora en que estaba programado para hacerlo".
Definiciones operacionales:	"Autorreporte de motivación intrínseca (cuestionario autoadministrado) del Inventario de Características del Trabajo, versión mexicana".	"Revisión de los registros electrónicos de asistencia al trabajo durante el último trimestre".

El cuestionario de motivación intrínseca se desarrollaría y adaptaría al contexto del estudio en la fase del proceso cuantitativo denominada recolección de los datos; lo mismo ocurriría con el procedimiento para medir el "ausentismo laboral". Desde luego, también durante esta etapa las variables llegan a ser objeto de modificación o ajuste y, en consecuencia, también sus definiciones.

Resumen

- En este punto de la investigación es necesario analizar si es conveniente formular o no hipótesis, lo que depende del alcance inicial del estudio (exploratorio, descriptivo, correlacional o explicativo).
- Las hipótesis son proposiciones tentativas acerca de las relaciones entre dos o más variables y se apoyan en conocimientos organizados y sistematizados.
- Las hipótesis son el centro del enfoque cuantitativo-deductivo.
- Las hipótesis contienen variables; éstas son propiedades cuya variación es susceptible de ser medida, observada o inferida.
- Las hipótesis surgen normalmente del planteamiento del problema y la revisión de la literatura, y algunas veces a partir de teorías.
- Las hipótesis deben referirse a una situación, un contexto, un ambiente o un evento empírico. Las variables contenidas deben ser precisas, concretas, y poder observarse en la realidad; la relación entre las variables debe ser clara, verosímil y medible. Asimismo, las hipótesis tienen que vincularse con técnicas disponibles para probarlas.
- Al definir el alcance del estudio (exploratorio, descriptivo, correlacional o explicativo), el investigador decide formular o no hipótesis. En los estudios exploratorios no se establecen hipótesis.
- Las hipótesis se clasifican en: *a*) hipótesis de investigación, *b*) hipótesis nulas, *c*) hipótesis alternativas y *d*) hipótesis estadísticas.
- A su vez, las hipótesis de investigación se clasifican de la manera como se muestra en la figura 6.7.
- Puesto que las hipótesis nulas y las alternativas se derivan de las hipótesis de investigación, pueden clasificarse del mismo modo, pero con los elementos que las caracterizan.
- Centro de recursos en línea Las hipótesis estadísticas se clasifican en: *a*) hipótesis estadísticas de estimación, *b*) hipótesis estadísticas de correlación y *c*) hipótesis estadísticas de la diferencia de grupos. Son propias de estudios cuantitativos. Éstas se revisan en el capítulo 8 del Centro de Recursos en Línea: "Análisis estadístico: segunda parte", el cual se encuentra en Material complementario → Capítulos, y se puede descargar.
- En una investigación pueden formularse una o varias hipótesis de distintos tipos.
- Dentro del enfoque deductivo-cuantitativo, las hipótesis se contrastan con la realidad para aceptarse o rechazarse en un contexto determinado.
- Las hipótesis constituyen las guías de una investigación.
- La formulación de hipótesis va acompañada de las definiciones conceptuales y operacionales de las variables contenidas dentro de la hipótesis.
- Una definición conceptual trata a la variable con otros términos, es como una definición de diccionario especializado.
- La definición operacional indica cómo vamos a medir la variable.
- Hay investigaciones en las que no se puede formular hipótesis porque el fenómeno que se va a estudiar es desconocido o se carece de información para establecerlas (pero sólo ocurre en los estudios exploratorios y algunos estudios descriptivos).

Figura 6.7 Clasificación de las hipótesis de investigación.

- ***a)*** Hipótesis descriptiva de un dato o valor que se pronostica
- ***b)*** Hipótesis correlacionales
  - Hipótesis que establecen simplemente relación entre las variables
    - Bivariadas
    - Multivariadas
  - Hipótesis que establecen cómo es la relación entre las variables (hipótesis direccionales)
    - Bivariadas
    - Multivariadas
- ***c)*** Hipótesis de la diferencia de grupos
  - Hipótesis que establecen diferencias entre los grupos a comparar
  - Hipótesis que especifican en favor de qué grupo (de los que se comparan) es la diferencia
- ***d)*** Hipótesis causales
  - Bivariadas
  - Multivariadas
    - Hipótesis con varias variables independientes y una dependiente
    - Hipótesis con una variable independiente y varias dependientes
    - Hipótesis con diversas variables tanto independientes como dependientes
    - Hipótesis con presencia de variables intervinientes
    - Hipótesis altamente complejas

5.3. Metodología

5.3.1. Caracterización de la investigación: enfoque, alcance y tipo (Diseño)

- Autor principal: **Roberto Hernández Sampieri**
- Libro recomendado: *Metodología de la Investigación*

5.3.2. Selección de técnicas e instrumentos para la recolección de datos

- Autor principal: **Fidias G. Arias**
- Libro recomendado: *El Proyecto de Investigación*
- Autor de apoyo: **Roberto Hernández Sampieri**
- Obra: *Metodología de la Investigación*

5.3.1. Caracterización de la investigación: enfoque, alcance y tipo (Diseño)

Para el desarrollo del subtema **Caracterización de la investigación: enfoque, alcance y tipo**, se consideran aquellos aspectos que permiten determinar la orientación metodológica de una investigación, a partir de la identificación del enfoque, el alcance y el diseño de investigación. Hernández Sampieri y otros autores presentan los enfoques cuantitativo y cualitativo, sus características y diferencias; asimismo, desarrollan los alcances exploratorio, descriptivo, correlacional y explicativo, además de la concepción o elección del diseño de investigación. Estos elementos permiten establecer las características metodológicas que orientarán el desarrollo del estudio.

N.°	Libros	Capítulo	Página	Contenido
1	*Metodología de la investigación*	**Capítulo 1. Definiciones de los enfoques cuantitativo y cualitativo, sus similitudes y diferencias**	**p. 4**	**¿Qué enfoques se han presentado en la investigación?** Presenta los enfoques que se han desarrollado en la investigación científica.
2	*Metodología de la investigación*	**Capítulo 1. Definiciones de los enfoques cuantitativo y cualitativo, sus similitudes y diferencias**	**p. 4**	**¿Qué características posee el enfoque cuantitativo de investigación?** Describe las características propias del enfoque cuantitativo.
3	*Metodología de la investigación*	**Capítulo 1. Definiciones de los enfoques cuantitativo y cualitativo, sus similitudes y diferencias**	**p. 7**	**¿Qué características posee el enfoque cualitativo de investigación?** Describe las características propias del enfoque cualitativo.
4	*Metodología de la investigación*	**Capítulo 1. Definiciones de los enfoques cuantitativo y cualitativo, sus similitudes y diferencias**	**p. 10**	**¿Cuáles son las diferencias entre los enfoques cuantitativo y cualitativo?** Establece las diferencias entre ambos enfoques de investigación.
5	*Metodología de la investigación*	**Capítulo 5. Definición del alcance de la investigación que se realizará: exploratorio, descriptivo, correlacional o explicativo**	**p. 90**	**¿Qué alcances puede tener el proceso de investigación cuantitativa?** Presenta los diferentes alcances que puede asumir una investigación cuantitativa.

6	*Metodología de la investigación*	**Capítulo 5. Definición del alcance de la investigación que se realizará: exploratorio, descriptivo, correlacional o explicativo**	**p. 91**	**¿En qué consisten los estudios de alcance exploratorio?** Explica las características y el valor de los estudios exploratorios.
7	*Metodología de la investigación*	**Capítulo 5. Definición del alcance de la investigación que se realizará: exploratorio, descriptivo, correlacional o explicativo**	**p. 92**	**¿En qué consisten los estudios de alcance descriptivo?** Explica las características y el valor de los estudios descriptivos.
8	*Metodología de la investigación*	**Capítulo 5. Definición del alcance de la investigación que se realizará: exploratorio, descriptivo, correlacional o explicativo**	**p. 93**	**¿En qué consisten los estudios de alcance correlacional?** Explica la utilidad, el valor y las características de los estudios correlacionales.
9	*Metodología de la investigación*	**Capítulo 5. Definición del alcance de la investigación que se realizará: exploratorio, descriptivo, correlacional o explicativo**	**p. 95**	**¿En qué consisten los estudios de alcance explicativo?** Explica las características de los estudios explicativos y su grado de estructuración.
10	*Metodología de la investigación*	**Capítulo 7. Concepción o elección del diseño de investigación**	**p. 126**	**Concepción o elección del diseño de investigación.** Introduce el diseño de investigación y su elección dentro del proceso cuantitativo.

¿Cómo se define la investigación?

La **investigación** es un conjunto de procesos sistemáticos, críticos y empíricos que se aplican al estudio de un fenómeno o problema.

¿Qué enfoques se han presentado en la investigación?

1

A lo largo de la historia de la ciencia han surgido diversas corrientes de pensamiento (como el empirismo, el materialismo dialéctico, el positivismo, la fenomenología, el estructuralismo) y diversos marcos interpretativos, como el realismo y el constructivismo, que han abierto diferentes rutas en la búsqueda del conocimiento. No se profundizará por ahora en ellas; su revisión, aunque breve, se incluye en el centro de recursos en línea, en: http://www.mhhe.com/he/hmi6e de esta edición.[1] Sin embargo, y debido a las diferentes premisas que las sustentan, desde el siglo pasado tales corrientes se "polarizaron" en dos aproximaciones principales de la investigación: el enfoque cuantitativo y el enfoque cualitativo.[2]

Centro de recursos en línea

Ambos enfoques emplean procesos cuidadosos, metódicos y empíricos en su esfuerzo para generar conocimiento, por lo que la definición previa de investigación se aplica a los dos por igual. En términos generales, estos métodos utilizan cinco estrategias similares y relacionadas entre sí (Grinnell, 1997):

1. Llevan a cabo la observación y evaluación de fenómenos.
2. Establecen suposiciones o ideas como consecuencia de la observación y evaluación realizadas.
3. Demuestran el grado en que las suposiciones o ideas tienen fundamento.
4. Revisan tales suposiciones o ideas sobre la base de las pruebas o del análisis.
5. Proponen nuevas observaciones y evaluaciones para esclarecer, modificar y fundamentar las suposiciones e ideas o incluso para generar otras.

Sin embargo, aunque las aproximaciones cuantitativa y cualitativa comparten esas estrategias generales, cada una tiene sus propias características.

¿Qué características posee el enfoque cuantitativo de investigación?

2

El **enfoque cuantitativo** (que representa, como dijimos, un conjunto de procesos) es secuencial y probatorio. Cada etapa precede a la siguiente y no podemos "brincar" o eludir pasos.[3] El orden es riguroso, aunque desde luego, podemos redefinir alguna fase. Parte de una idea que va acotándose y, una vez delimitada, se derivan objetivos y preguntas de investigación, se revisa la literatura y se construye un marco o una perspectiva teórica. De las preguntas se establecen hipótesis y determinan variables; se traza un plan para probarlas (diseño); se miden las variables en un determinado contexto; se analizan las mediciones obtenidas utilizando métodos estadísticos, y se extrae una serie de conclusiones

Enfoque cuantitativo Utiliza la recolección de datos para probar hipótesis con base en la medición numérica y el análisis estadístico, con el fin establecer pautas de comportamiento y probar teorías.

Centro de recursos en línea

[1] En el centro de recursos en línea: http://www.mhhe.com/he/hmi6e, el lector encontrará un capítulo sobre los antecedentes de las aproximaciones cuantitativa y cualitativa (véase el primer capítulo de los adicionales: "Historia de los enfoques cuantitativo, cualitativo y mixto").

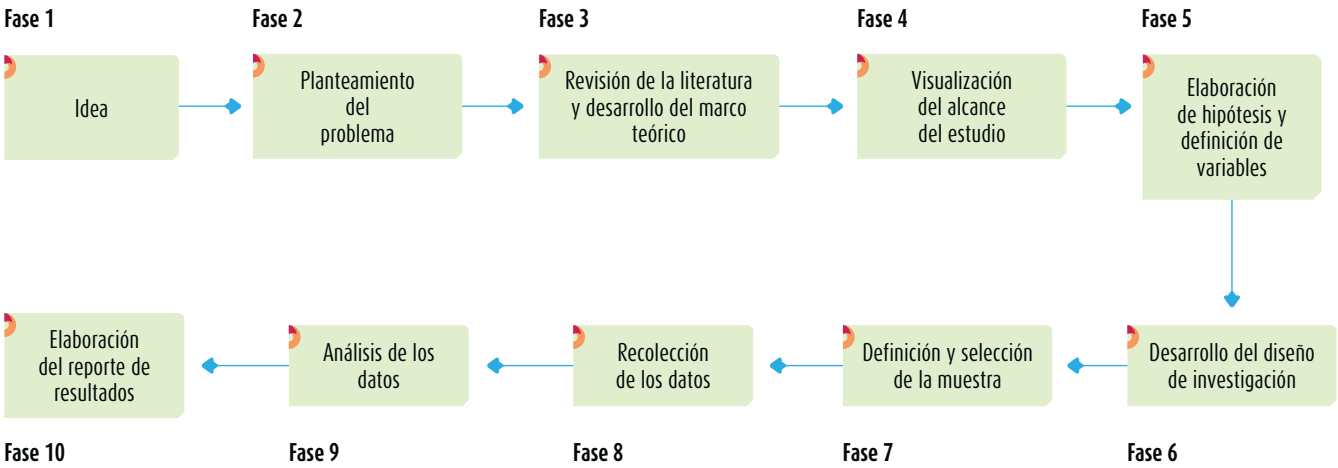
[2] El significado original del término "cuantitativo" (del latín *quantitas*) se remite a conteos numéricos y métodos matemáticos, mientras que la palabra "cualitativa" (del latín *qualitas*) hace referencia a la naturaleza, carácter y propiedades de los fenómenos (Niglas, 2010), y aunque en el centro de recursos en línea se profundiza más en este tema, por ahora basta decir que el enfoque cuantitativo en las ciencias sociales se origina fundamentalmente en la obra de Auguste Comte (1798-1857) y Émile Durkheim (1858-1917). Ellos propusieron que el estudio de los fenómenos sociales requiere ser "científico", es decir, susceptible a la aplicación del mismo método que se utilizaba con éxito en las ciencias naturales. Tales autores sostenían que todas las "cosas" o fenómenos que estudiaban las ciencias eran medibles. A esta corriente se le llama positivismo.

El enfoque cualitativo tiene su origen en otro pionero de las ciencias sociales: Max Weber (1864-1920), quien introdujo el término *verstehen* (en alemán, "entender"), reconociendo que además de la descripción y medición de variables sociales, deben considerarse los significados subjetivos y la comprensión del contexto en el que ocurre el fenómeno. Weber propuso un método híbrido, con herramientas como los tipos ideales, en el que los estudios no sean únicamente de variables macrosociales, sino de instancias individuales.

[3] Por ejemplo, no podemos definir y seleccionar la muestra si aún no hemos establecido las hipótesis. Tampoco es posible recolectar o analizar datos si previamente no hemos desarrollado el diseño o no hemos definido la muestra.

respecto de la o las hipótesis. Este proceso se representa en la figura 1.1 y se desarrollará en la segunda parte del libro.

Figura 1.1 Proceso cuantitativo.



El enfoque cuantitativo tiene las siguientes características: 3

1. Refleja la necesidad de medir y estimar magnitudes de los fenómenos o problemas de investigación: ¿cada cuánto ocurren y con qué magnitud?
2. El investigador o investigadora *plantea un problema de estudio delimitado y concreto* sobre el fenómeno, aunque en evolución. Sus preguntas de investigación versan sobre cuestiones específicas.
3. Una vez planteado el problema de estudio, el investigador o investigadora considera lo que se ha investigado anteriormente (la *revisión de la literatura*) y construye un *marco teórico* (la teoría que habrá de guiar su estudio), del cual deriva una o varias *hipótesis* (cuestiones que va a examinar si son ciertas o no) y las somete a prueba mediante el empleo de los diseños de investigación apropiados. Si los resultados corroboran las hipótesis o son congruentes con éstas, se aporta evidencia a su favor. Si se refutan, se descartan en busca de mejores explicaciones y nuevas hipótesis. Al apoyar las hipótesis se genera confianza en la teoría que las sustenta. Si *no* es así, se rechazan las hipótesis y, eventualmente, la teoría.
4. Así, las hipótesis (por ahora denominémoslas "creencias") se generan antes de recolectar y analizar los datos.
5. La *recolección de los datos* se fundamenta en la medición (se miden las variables o conceptos contenidos en las hipótesis). Esta recolección se lleva a cabo al utilizar procedimientos estandarizados y aceptados por una comunidad científica. Para que una investigación sea creíble y aceptada por otros investigadores, debe demostrarse que se siguieron tales procedimientos. Como en este enfoque se pretende *medir*, los fenómenos estudiados deben poder observarse o *referirse* al "mundo real".
6. Debido a que los datos son producto de mediciones, se representan mediante números (cantidades) y se deben *analizar* con *métodos estadísticos*.
7. En el proceso se trata de tener el mayor control para lograr que otras posibles explicaciones, distintas o "rivales" a la propuesta del estudio (hipótesis), se desechen y se excluya la incertidumbre y minimice el error. Es por esto que se confía en la experimentación o en las pruebas de causalidad.
8. Los análisis cuantitativos se interpretan a la luz de las predicciones iniciales (hipótesis) y de estudios previos (teoría). La interpretación constituye una explicación de cómo los resultados encajan en el conocimiento existente (Creswell, 2013a).

9. La investigación cuantitativa debe ser lo más "objetiva" posible.[4] Los fenómenos que se observan o miden no deben ser afectados por el investigador, quien debe evitar en lo posible que sus temores, creencias, deseos y tendencias influyan en los resultados del estudio o interfieran en los procesos y que tampoco sean alterados por las tendencias de otros (Unrau, Grinnell y Williams, 2005).
10. Los estudios cuantitativos siguen un patrón predecible y estructurado (el proceso) y se debe tener presente que las decisiones críticas sobre el método se toman antes de recolectar los datos.
11. En una investigación cuantitativa se intenta generalizar los resultados encontrados en un grupo o segmento (muestra) a una colectividad mayor (universo o población). También se busca que los estudios efectuados puedan replicarse.
12. Al final, con los estudios cuantitativos se pretende confirmar y predecir los fenómenos investigados, buscando regularidades y relaciones causales entre elementos. Esto significa que la meta principal es la formulación y demostración de teorías.
13. Para este enfoque, si se sigue rigurosamente el proceso y, de acuerdo con ciertas reglas lógicas, los datos generados poseen los estándares de validez y confiabilidad, las conclusiones derivadas contribuirán a la generación de conocimiento.
14. Esta aproximación se vale de la lógica o razonamiento deductivo, que comienza con la teoría, y de ésta se derivan expresiones lógicas denominadas "hipótesis" que el investigador somete a prueba.
15. La investigación cuantitativa pretende identificar leyes "universales" y causales (Bergman, 2008).
16. La búsqueda cuantitativa ocurre en la "realidad externa" al individuo. Conviene ahora explicar cómo se concibe la realidad con esta aproximación a la investigación.

Para este fin, utilizaremos la siguiente argumentación basada en Grinnell (1997) y Creswell (2013a):

1. Hay dos realidades: la primera es *interna* y consiste en las creencias, presuposiciones y experiencias *subjetivas* de las personas, que van desde las muy vagas o generales (intuiciones) hasta las convicciones bien organizadas y desarrolladas lógicamente a través de teorías formales. La segunda realidad es *objetiva*, *externa* e *independiente* de las creencias que tengamos sobre ella (la autoestima, una ley, los mensajes televisivos, una edificación, el sida, etc., ocurren, es decir, cada una constituye una realidad a pesar de lo que pensemos de ella).
2. Esta realidad objetiva es susceptible de conocerse. Bajo esta premisa, resulta posible investigar una realidad externa y autónoma del investigador.
3. Se necesita comprender o tener la mayor cantidad de información sobre la realidad objetiva. Conocemos la realidad del fenómeno y los eventos que la rodean a través de sus manifestaciones. Para entender cada realidad (el porqué de las cosas), es necesario registrar y analizar dichos eventos. Desde luego, en el *enfoque cuantitativo* lo subjetivo existe y posee un valor para los investigadores; pero este enfoque se aboca a demostrar qué tan bien se adecua el conocimiento a la realidad objetiva. Documentar esta coincidencia constituye un propósito central de muchos estudios cuantitativos (que los efectos que consideramos que provoca una enfermedad sean "verdaderos", que captemos la relación "real" entre las motivaciones de una persona y su conducta, que un material que se supone posea determinada resistencia auténticamente la tenga, entre otros).
4. Cuando las investigaciones creíbles establezcan que la *realidad objetiva* es diferente de nuestras creencias, éstas deben modificarse para adaptarse a tal realidad, como se ve en la figura 1.2 (note el lector que la "realidad" no cambia, sino que es la misma; lo que se ajusta es el conjunto de creencias o hipótesis del investigador y, en consecuencia, la teoría).

En el caso de las ciencias sociales, el enfoque cuantitativo parte de que el mundo "social" es intrínsecamente cognoscible y todos podemos estar de acuerdo con la naturaleza de la realidad social.

[4] Desde luego, sabemos que no existe la objetividad "pura o completa".

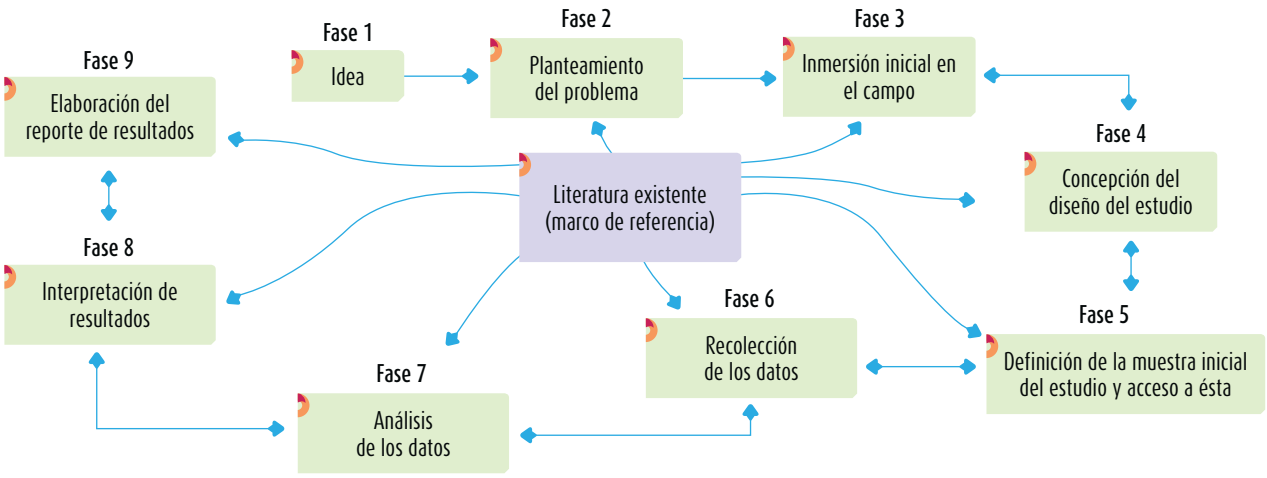
Figura 1.2 Relación entre la teoría, la investigación y la realidad en el enfoque cuantitativo.

¿Qué características posee el enfoque cualitativo de investigación?

El **enfoque cualitativo**[5] también se guía por áreas o temas significativos de investigación. Sin embargo, en lugar de que la claridad sobre las preguntas de investigación e hipótesis preceda a la recolección y el análisis de los datos (como en la mayoría de los estudios cuantitativos), los *estudios cualitativos* pueden desarrollar preguntas e hipótesis antes, durante o después de la recolección y el análisis de los datos. Con frecuencia, estas actividades sirven, primero, para descubrir cuáles son las preguntas de investigación más importantes; y después, para perfeccionarlas y responderlas. La acción indagatoria se mueve de manera dinámica en ambos sentidos: entre los hechos y su interpretación, y resulta un proceso más bien "circular" en el que la secuencia no siempre es la misma, pues varía con cada estudio. A continuación intentamos representarlo en la figura 1.3, pero cabe señalar que es simplemente eso, un intento, porque su complejidad y flexibilidad son mayores. Este proceso se despliega en la parte 3 del libro.

Enfoque cualitativo Utiliza la recolección y análisis de los datos para afinar las preguntas de investigación o revelar nuevas interrogantes en el proceso de interpretación.

Figura 1.3 Proceso cualitativo.



[5] Este enfoque también se conoce como investigación naturalista, fenomenológica, interpretativa o etnográfica, y es una especie de "paraguas" en el cual se incluye una variedad de concepciones, visiones, técnicas y estudios no cuantitativos. Según Sparkes y Smith (2014) y Savin-Baden y Major (2013), existen diversos marcos interpretativos, como el interaccionismo, la etnometodología, el constructivismo, el feminismo, la fenomenología, la psicología de los constructos personales, la teoría crítica, etc., que se incluyen en este "paraguas para efectuar estudios".

3 Para comprender la figura 1.3 es necesario observar lo siguiente:

a) Aunque ciertamente hay una revisión inicial de la literatura, ésta puede complementarse en cualquier etapa del estudio y apoyar desde el planteamiento del problema hasta la elaboración del reporte de resultados (la vinculación entre la teoría y las etapas del proceso se representa mediante flechas curvadas).

b) En la investigación cualitativa a veces es necesario regresar a etapas previas. Por ello, las flechas de las fases que van de la inmersión inicial en el campo hasta el reporte de resultados se visualizan en dos sentidos.

Por ejemplo, el primer diseño del estudio puede modificarse al definir la muestra inicial y pretender tener acceso a ésta (podría ser el caso que se desee observar a ciertas personas en su ambiente natural, pero por alguna razón descubrimos que no es factible efectuar las observaciones deseadas; en consecuencia, la muestra y los ambientes de estudio tienen que variar, y el diseño debe ajustarse). Tal fue la situación de un estudiante que deseaba observar en un penal a criminales de alta peligrosidad con ciertas características, pero le fue negado el acceso y tuvo que acudir a otra cárcel, donde entrevistó a criminales menos peligrosos.

Asimismo, al analizar los datos, podemos advertir que necesitamos un número mayor de participantes u otras personas que al principio no estaban contempladas, lo cual modifica la muestra concebida originalmente. O bien, descubrimos que debemos analizar otra clase de datos no considerados al inicio del estudio (por ejemplo, habíamos planeado efectuar únicamente entrevistas y nos encontramos con documentos valiosos de los individuos que nos pueden ayudar a comprenderlos mejor, como sería el caso de sus "diarios personales").

c) La inmersión inicial en el campo significa sensibilizarse con el ambiente o entorno en el cual se llevará a cabo el estudio, identificar informantes que aporten datos y guíen al investigador por el lugar, adentrarse y compenetrarse con la situación de investigación, además de verificar la factibilidad del estudio.

d) En el caso del proceso cualitativo, la muestra, la recolección y el análisis son fases que se realizan prácticamente de manera simultánea.

2 Además de lo anterior, el *enfoque o aproximación cualitativa* posee las siguientes características:

1. El investigador o investigadora plantea un problema, pero no sigue un proceso definido claramente. Sus planteamientos iniciales *no* son tan específicos como en el enfoque cuantitativo y las preguntas de investigación *no* siempre se han conceptualizado ni definido por completo.
2. En la búsqueda cualitativa, en lugar de iniciar con una teoría y luego "voltear" al mundo empírico para confirmar si ésta es apoyada por los datos y resultados, el investigador comienza examinando los hechos en sí y en el proceso desarrolla una teoría coherente para representar lo que observa (Esterberg, 2002). Dicho de otra forma, las *investigaciones cualitativas* se basan más en una lógica y proceso inductivo (explorar y describir, y luego generar perspectivas teóricas). Van de lo particular a lo general. Por ejemplo, en un estudio cualitativo típico, el investigador entrevista a una persona, analiza los datos que obtuvo y saca conclusiones; posteriormente, entrevista a otra persona, analiza esta nueva información y revisa sus resultados y conclusiones; del mismo modo, efectúa y analiza más entrevistas para comprender el fenómeno que estudia. Es decir, procede caso por caso, dato por dato, hasta llegar a una perspectiva más general.
3. En la mayoría de los estudios cualitativos no se prueban hipótesis, sino que se generan durante el proceso y se perfeccionan conforme se recaban más datos; son un resultado del estudio.
4. El enfoque se basa en métodos de recolección de datos *no* estandarizados ni predeterminados completamente. Tal recolección consiste en obtener las perspectivas y puntos de vista de los participantes (sus emociones, prioridades, experiencias, significados y otros aspectos más bien subjetivos). También resultan de interés las interacciones entre individuos, grupos y colectividades. El investigador hace preguntas más abiertas, recaba datos expresados a través del lenguaje escrito, verbal y no verbal, así como visual, los cuales describe, analiza y convierte en temas que vincula, y reconoce sus tendencias personales. Debido a ello, la preocupación directa del investigador se concentra en las vivencias de los participantes tal como fueron (o son) sentidas y experimentadas

(Sherman y Webb, 1988). Patton (2011) define los **datos cualitativos** como descripciones detalladas de situaciones, eventos, personas, interacciones, conductas observadas y sus manifestaciones.

Datos cualitativos Evidencia o información simbólica verbal, audiovisual o en forma de texto e imágenes.

5. Así, el investigador cualitativo utiliza técnicas para recolectar datos, como la observación no estructurada, entrevistas abiertas, revisión de documentos, discusión en grupo, evaluación de experiencias personales, registro de historias de vida, e interacción e introspección con grupos o comunidades.
6. El proceso de indagación es más flexible y se mueve entre las respuestas y el desarrollo de la teoría. Su propósito consiste en "reconstruir" la realidad, tal como la observan los actores de un sistema social definido previamente. Es *holístico*, porque se precia de considerar el "todo"[6] sin reducirlo al estudio de sus partes.
7. La aproximación cualitativa evalúa el desarrollo natural de los sucesos, es decir, no hay manipulación ni estimulación de la realidad (Corbetta, 2003).
8. La investigación cualitativa se fundamenta en una perspectiva interpretativa centrada en el entendimiento del significado de las acciones de seres vivos, sobre todo de los humanos y sus instituciones (busca interpretar lo que va captando activamente).
9. Postula que la "realidad" se define a través de las interpretaciones de los participantes en la investigación respecto de sus propias realidades. De este modo, convergen varias "realidades", por lo menos la de los participantes, la del investigador y la que se produce en la interacción de todos los actores. Además, son realidades que van modificándose conforme transcurre el estudio y son las fuentes de datos.
10. Por lo anterior, el investigador se introduce en las experiencias de los participantes y construye el conocimiento, siempre consciente de que es parte del fenómeno estudiado. Así, en el centro de la investigación está situada la diversidad de ideologías y cualidades únicas de los individuos.
11. Las indagaciones cualitativas no pretenden generalizar de manera probabilística los resultados a poblaciones más amplias ni obtener necesariamente muestras representativas; incluso, regularmente no pretenden que sus estudios lleguen a repetirse.
12. El enfoque cualitativo puede concebirse como un conjunto de prácticas interpretativas que hacen al mundo "visible", lo transforman y convierten en una serie de representaciones en forma de observaciones, anotaciones, grabaciones y documentos. Es *naturalista* (porque estudia los fenómenos y seres vivos en sus contextos o ambientes naturales y en su cotidianidad) e *interpretativo* (pues intenta encontrar sentido a los fenómenos en función de los significados que las personas les otorguen).

En la aproximación cualitativa hay una variedad de concepciones o marcos de interpretación, que guardan un común denominador: todo individuo, grupo o sistema social tiene una manera única de ver el mundo y entender situaciones y eventos, la cual se construye por el inconsciente, lo transmitido por otros y por la experiencia, y mediante la investigación, debemos tratar de comprenderla en su contexto.

Creswell (2013b) y Neuman (1994) sintetizan las actividades principales del investigador o investigadora cualitativa con los siguientes comentarios:

- Adquiere un punto de vista "interno" (desde dentro del fenómeno), aunque mantiene una perspectiva analítica o cierta distancia como observador externo.
- Utiliza diversas técnicas de investigación y habilidades sociales de una manera flexible, de acuerdo con los requerimientos de la situación.
- No define las variables con el propósito de manipularlas experimentalmente.
- Produce datos en forma de notas extensas, diagramas, mapas o "cuadros humanos" para generar descripciones bastante detalladas.

[6] Aquí el "todo" es el fenómeno de interés. Por ejemplo, en su libro *Police Work*, Peter Manning (1997) se sumerge por semanas en el estudio y análisis del trabajo policiaco. Le interesa comprender las relaciones y lealtades que surgen entre personas que se dedican a esta profesión. Lo logra sin "medición" de actitudes, tan sólo captando el fenómeno mismo de la vida en la policía.

- Extrae significado de los datos y no necesita reducirlos a números ni debe analizarlos estadísticamente (aunque el conteo puede utilizarse en el análisis).
- Entiende a los participantes que son estudiados y desarrolla una empatía hacia ellos; no sólo registra hechos "objetivos".
- Mantiene una doble perspectiva: analiza los aspectos explícitos, conscientes y manifiestos, así como aquellos implícitos, inconscientes y subyacentes. En este sentido, la realidad subjetiva en sí misma es objeto de estudio.
- Observa los procesos sin irrumpir, alterar ni imponer un punto de vista externo, sino tal como los perciben los actores del sistema social.
- Es capaz de manejar paradojas, incertidumbres, dilemas éticos y ambigüedades.

¿Cuáles son las diferencias entre los enfoques cuantitativo y cualitativo?

El *enfoque cualitativo* busca principalmente la "dispersión o expansión" de los datos e información, mientras que el *enfoque cuantitativo* pretende "acotar" intencionalmente la información (medir con precisión las variables del estudio, tener "foco").[7]

4

Mientras que un estudio cuantitativo se basa en investigaciones previas, el estudio cualitativo se fundamenta primordialmente en sí mismo. El cuantitativo se utiliza para consolidar las creencias (formuladas de manera lógica en una teoría o un esquema teórico) y establecer con exactitud patrones de comportamiento de una población; y el cualitativo, para que el investigador se forme creencias propias sobre el fenómeno estudiado, como lo sería un grupo de personas únicas o un proceso particular.

Para reforzar las características de ambas aproximaciones y ahondar en sus diferencias, hemos preferido compararlas en la tabla 1.1. Algunas concepciones han sido adaptadas o reformuladas de diversos autores.[8]

Tabla 1.1 Diferencias entre los enfoques cuantitativo y cualitativo

Definiciones (dimensiones)	Enfoque cuantitativo	Enfoque cualitativo
Marcos generales de referencia básicos	Positivismo, neopositivismo y pospositivismo.	Fenomenología, constructivismo, naturalismo, interpretativismo.
Punto de partida*	Hay una realidad que conocer. Esto puede hacerse a través de la mente.	Hay una realidad que descubrir, construir e interpretar. La realidad es la mente.
Realidad que se va a estudiar	Existe una realidad objetiva única. El mundo es concebido como externo al investigador.	Existen varias realidades subjetivas construidas en la investigación, las cuales varían en su forma y contenido entre individuos, grupos y culturas. Por ello, el investigador cualitativo parte de la premisa de que el mundo social es "relativo" y sólo puede ser entendido desde el punto de vista de los actores estudiados.

* Becker (1993) dice: la "realidad" es el punto más estresante en las ciencias sociales. Las diferencias entre los dos enfoques han tenido un tinte eminentemente ideológico. El gran filósofo alemán Karl Popper (1965) explica que las visiones conflictivas sobre lo que es o debe ser el estudio del fenómeno social se originan en las premisas de diferentes definiciones de lo que es la realidad. El realismo, desde Aristóteles, establece que el mundo llega a ser conocido por la mente. Kant introduce que el mundo puede ser conocido porque la realidad se asemeja a las formas que la mente tiene. En tanto que Hegel va hacia un idealismo puro y propone: "El mundo es mi mente". Esto último es ciertamente confuso, y así lo considera Popper, advirtiendo que el gran peligro de esta posición es que permite el dogmatismo (como lo ha probado con el ejemplo del materialismo dialéctico). El avance en el conocimiento, dice Popper, necesita de conceptos que podamos refutar o probar. Esta característica delimita qué es y qué no es ciencia.

(*continúa*)

[7] Usemos el ejemplo de una cámara fotográfica: en el estudio *cuantitativo* se define lo que se va a fotografiar y se toma la foto. En el *cualitativo* es como si la función de "*zoom in*" (acercamiento) y "*zoom out*" (alejamiento) se utilizaran constantemente para capturar en un área cualquier figura de interés.

[8] Miller-Cochran y Rodrigo (2014), Peters (2014), Hernández-Sampieri, Zapata y Mendoza (2013); Creswell (2013a), Morgan (2013), Hartas (2013), Shaffer (2013), Maxwell (2013), Pratt (2013), Babbie (2012), Creswell y Plano Clark (2011), Grinnell y Unrau (2011), Morris y Burkett (2011), Yin (2011), Hesee-Biber (2010), Onwuegbuzie y Combs (2010), Stake (2010), Teddlie y Tashakkori (2009), Corbin y Strauss (2007), Todd y Lobeck (2004), Corbetta (2003), Sandín (2003), Álvarez-Gayou (2003), Esterberg (2002), Guba y Lincoln (1994).

www.elosopanda.com | jamespoetrodriguez.com

Tabla 1.1 (*continuación*)

Definiciones (dimensiones)	Enfoque cuantitativo	Enfoque cualitativo
Naturaleza de la realidad	La realidad no cambia por las observaciones y mediciones realizadas.*	La realidad sí cambia por las observaciones y la recolección de datos.
Objetividad	Busca ser objetivo.	Admite subjetividad.
Metas de la investigación	Describir, explicar, comprobar y predecir los fenómenos (causalidad). Generar y probar teorías.	Describir, comprender e interpretar los fenómenos, a través de las percepciones y significados producidos por las experiencias de los participantes.
Lógica	Se aplica la lógica deductiva. De lo general a lo particular (de las leyes y teoría a los datos).	Se aplica la lógica inductiva. De lo particular a lo general (de los datos a las generalizaciones —no estadísticas— y la teoría).
Relación entre ciencias físicas/naturales y sociales	Las ciencias físicas/naturales y las sociales son una unidad. A las ciencias sociales pueden aplicárseles los principios de las ciencias naturales.	Las ciencias físicas/naturales y las sociales son diferentes. No se aplican los mismos principios.
Posición personal del investigador	Neutral. El investigador "hace a un lado" sus propios valores y creencias. La posición del investigador es "imparcial", intenta asegurar procedimientos rigurosos y "objetivos" de recolección y análisis de los datos, así como evitar que sus sesgos y tendencias influyan en los resultados.	Explícita. El investigador reconoce sus propios valores y creencias, incluso son fuentes de datos parte del estudio.
Papel de los fenómenos estudiados (objetos, seres vivos, etcétera)	Los papeles son más bien pasivos.	Los papeles son más bien activos (incluso los objetos pueden poseer diferentes "significados").
Relación básica entre el investigador y el fenómeno estudiado	De independencia y neutralidad, no se afectan. Se separan.	De interdependencia, se influyen. No se separan.
Uso de la teoría	La teoría se utiliza para ajustar sus postulados al "mundo empírico".	La teoría es un marco de referencia.
Generación de la teoría	La teoría se genera a partir de comparar la investigación previa con los resultados del estudio. De hecho, éstos son una extensión de las investigaciones antecedentes.	La teoría se construye básicamente a partir de los datos empíricos obtenidos y analizados y, desde luego, se compara con los resultados de estudios anteriores.
Papel de la revisión de la literatura	La literatura representa un papel crucial, guía a la investigación. Es fundamental para el planteamiento, la definición de la teoría, las hipótesis, el diseño y demás etapas del proceso.	La literatura desempeña un papel menos importante al inicio, aunque sí es relevante en el desarrollo del proceso. En ocasiones, provee de dirección, pero lo que principalmente señala el rumbo es la evolución de eventos durante el estudio y el aprendizaje que se obtiene de los participantes.
La revisión de la literatura y las variables o conceptos de estudio	El investigador hace una revisión de la literatura principalmente para buscar variables significativas que puedan ser medidas.	El investigador, más que fundamentarse en la revisión de la literatura para seleccionar y definir las variables o conceptos clave del estudio, confía en el proceso mismo de investigación para identificarlos y descubrir cómo se relacionan.
Hipótesis	Se prueban hipótesis. Éstas se establecen para aceptarlas o rechazarlas dependiendo del grado de certeza (probabilidad).	Se generan hipótesis durante el estudio o al final de éste.
Diseño de la investigación	Estructurado, predeterminado (precede a la recolección de los datos).	Abierto, flexible, construido durante el trabajo de campo o realización del estudio.

(*continúa*)

*Aunque algunos físicos al estudiar las partículas se han percatado de lo relativo que resulta esta aseveración.

Tabla 1.1 (*continuación*)

Definiciones (dimensiones)	Enfoque cuantitativo	Enfoque cualitativo
Población-muestra	El objetivo es generalizar los datos de una muestra a una población (de un grupo pequeño a uno mayor).	Regularmente no se pretende generalizar los resultados obtenidos en la muestra a una población.
Muestra	Se involucran a muchos casos en la investigación porque se pretende generalizar los resultados del estudio.	Se involucran a unos cuantos casos porque no se pretende necesariamente generalizar los resultados del estudio, sino analizarlos intensivamente.
Composición de la muestra	Casos que en conjunto son estadísticamente representativos.	Casos individuales, representativos no desde el punto de vista estadístico, sino por sus "cualidades".
Naturaleza de los datos	La naturaleza de los datos es cuantitativa (datos numéricos).	La naturaleza de los datos es cualitativa (textos, narraciones, significados, etcétera).
Tipo de datos	Datos confiables y duros. En inglés: *hard*.	Datos profundos y enriquecedores. En inglés: *soft*.
Recolección de los datos	La recolección se basa en instrumentos estandarizados. Es uniforme para todos los casos. Los datos se obtienen por observación, medición y documentación. Se utilizan instrumentos que han demostrado ser válidos y confiables en estudios previos o se generan nuevos basados en la revisión de la literatura y se prueban y ajustan. Las preguntas, ítems o indicadores utilizados son específicos con posibilidades de respuesta o categorías predeterminadas.	La recolección de los datos está orientada a proveer de un mayor entendimiento de los significados y experiencias de las personas. El investigador es el instrumento de recolección de los datos, se auxilia de diversas técnicas que se desarrollan durante el estudio. Es decir, no se inicia la recolección de los datos con instrumentos preestablecidos, sino que el investigador comienza a aprender por observación y descripciones de los participantes y concibe formas para registrar los datos que se van refinando conforme avanza la investigación.
Concepción de los participantes en la recolección de datos	Los participantes son fuentes externas de datos.	Los participantes son fuentes internas de datos. El investigador también es un participante.
Finalidad del análisis de los datos	Describir las variables y explicar sus cambios y movimientos.	Comprender a las personas, procesos, eventos y sus contextos.
Características del análisis de los datos	• Sistemático y estandarizado. Utilización intensiva de la estadística (descriptiva e inferencial). • Basado en variables y casos (una matriz). • Impersonal. • Posterior a la recolección de los datos.	• Progresivo y varía dependiendo del modo en que se recolecten los datos y el tipo de éstos (textuales, simbólicos, en audio y/o video). • Fundamentado en la inducción analítica. • Uso moderado de la estadística (conteo, algunas operaciones aritméticas). • Basado en casos o personas y sus manifestaciones. • El análisis consiste en describir información y desarrollar temas.
Forma de los datos para analizar	Los datos son representados en forma de números que son analizados estadísticamente.	Datos en forma de textos, imágenes, piezas audiovisuales, documentos y objetos personales.
Proceso del análisis de los datos	El análisis se inicia con ideas preconcebidas, basadas en las hipótesis formuladas. Una vez recolectados los datos numéricos, éstos se transfieren a una matriz, la cual se analiza mediante procedimientos estadísticos.	Por lo general, el análisis no se inicia con ideas preconcebidas sobre cómo se relacionan los conceptos o variables. Conforme se van reuniendo los datos verbales, en texto y/o audiovisuales, se integran en una base de datos, la cual se analiza para determinar significados y describir el fenómeno estudiado desde el punto de vista de sus actores. Se conjuntan descripciones de participantes con las del investigador.

(*continúa*)

Tabla 1.1 (*continuación*)

Definiciones (dimensiones)	Enfoque cuantitativo	Enfoque cualitativo
Perspectiva del investigador en el análisis de los datos	Externa (al margen de los datos). El investigador no involucra sus creencias ni tendencias en el análisis. Mantiene distancia de éste.	Interna (desde los datos). El investigador involucra en el análisis sus propias creencias, así como la relación que tuvo con los participantes del estudio.
Principales criterios de evaluación en la recolección y análisis de los datos	Objetividad, rigor, confiabilidad y validez.	Credibilidad, confirmación, valoración, representatividad de voces y transferencia.
Presentación de resultados	Tablas, diagramas y modelos estadísticos. El formato de presentación es relativamente estándar.	El investigador emplea una variedad de formatos para reportar sus resultados: narraciones, fragmentos de textos, videos, audios, fotografías y mapas; diagramas, matrices y modelos conceptuales.
Reporte de resultados	Los reportes utilizan un tono objetivo, impersonal, no emotivo.	Los reportes utilizan un tono personal y emotivo.

Con el propósito de que el lector que se inicia en estos menesteres tenga una idea de la diferencia entre ambas aproximaciones, utilizaremos un ejemplo muy sencillo y cotidiano relativo a la atracción física, aunque a algunas personas podría parecerles simple. Desde luego, en el ejemplo no se consideran las implicaciones paradigmáticas que se encuentran detrás de cada enfoque, pero sí se hace hincapié en que, en términos prácticos, ambos contribuyen al conocimiento de un fenómeno.

Ejemplo

Comprensión de los enfoques cuantitativo y cualitativo de la investigación

Supongamos que un estudiante se encuentra interesado en saber qué factores intervienen para que una persona sea definida y percibida como "atractiva y conquistadora" (que cautiva a individuos del género opuesto y logra que se sientan atraídos hacia él o ella y se enamoren). Entonces, decide llevar a cabo un estudio (su idea para investigar) en su escuela.

Según el enfoque cuantitativo-deductivo, el estudiante plantearía su problema de investigación definiendo su objetivo y su pregunta (lo que quiere hacer y lo que quiere saber).

Por ejemplo, el objetivo podría ser: "conocer los factores que determinan que una persona joven sea percibida como atractiva y conquistadora", y la pregunta de investigación: "¿qué factores determinan que una persona joven sea percibida como atractiva y conquistadora?"

Después, revisaría estudios sobre la atracción física y psicológica en las relaciones entre jóvenes, la percepción de los jóvenes en torno a dichas relaciones, los elementos que intervienen en el inicio de la convivencia amorosa, las diferencias por género de acuerdo con los atributos y cualidades que les atraen de los demás, etcétera.

Precisaría su problema de investigación. Seleccionaría una teoría que explicara de manera satisfactoria —sobre la base de estudios previos— la atracción física y psicológica, la percepción de atributos y cualidades deseables en personas del género opuesto y el enamoramiento en las relaciones entre jóvenes; asimismo, y de ser posible, establecería una o varias hipótesis. Por ejemplo: "los chicos y las chicas que logran más conquistas amorosas y son percibidos como más 'atractivos' resultan ser aquellos que tienen mayor prestigio social en la escuela, que son más seguros de sí mismos y más extravertidos".

Después, podría entrevistar a compañeras y compañeros de su escuela y los interrogaría sobre el grado en que el prestigio social, la seguridad en uno mismo y la extraversión influyen en la "conquista" y "el atractivo" hacia personas del otro género. Incluso llegaría a utilizar cuestionarios ya establecidos, bien diseñados y confiables. Tal vez entrevistaría sólo a una muestra de estudiantes. También sería posible preguntar a las personas jóvenes que tienen fama de conquistadoras y atractivas qué piensan al respecto.

Además, analizaría los datos y la información producto de las entrevistas para obtener conclusiones acerca de sus hipótesis. Quizá también experimentaría eligiendo a individuos jóvenes que tuvieran diferentes grados de prestigio, seguridad y extraversión (niveles del perfil "conquistador y atractivo"), lanzándolos a conquistar a jóvenes del género opuesto para evaluar los resultados.

Su interés sería generalizar sus descubrimientos, al menos en relación con lo que ocurre en su comunidad estudiantil. Busca probar sus creencias, y si resulta que *no* consigue demostrar que el prestigio, la seguridad en sí mismo y la extraversión son factores relacionados con la conquista y el atractivo, intentaría otras explicaciones. Tal vez agregaría factores como la manera en que se visten, si son cosmopolitas (si han viajado mucho,

Un tema de la investigación cuantitativa-deductiva podría ser el siguiente: "¿Qué factores determinan que una persona joven sea percibida como atractiva y conquistadora?

conocen otras culturas), la inteligencia emocional y otros aspectos.

En el proceso irá deduciendo de la teoría lo que encuentra en su estudio. Desde luego, si la teoría que seleccionó es inadecuada, sus resultados serán insuficientes.

Con el enfoque cualitativo-inductivo, más que comenzar revisando las teorías sobre ciertos factores, lo que haría el estudiante sería sentarse en la cafetería a observar a chicos y chicas que tienen fama de ser atractivos y conquistadores. Observaría a la primera persona joven que considere que tiene esas características, la analizaría y formaría un concepto de ella (¿cómo es?, ¿qué perfil tiene?, ¿cómo se comporta?, ¿cuáles son sus atributos y cualidades?, ¿de qué forma se relaciona con los demás?). Asimismo, sería testigo de cómo conquista a compañeros. Así, llegaría a algunas conclusiones. Posteriormente haría lo mismo (observar) con otras personas jóvenes. Poco a poco entendería por qué son percibidos esos compañeros como atractivos y conquistadores. De ahí, podría derivar algún esquema que explique las razones por las cuales estas personas conquistan a otras.

Después entrevistaría, por medio de preguntas abiertas, a estudiantes de ambos géneros (percibidos como atractivos) y también a quienes han sido conquistados por ellos. De ahí, de nueva cuenta, derivaría hallazgos y sacaría conclusiones, y podría fundamentar algunas hipótesis, que al final contrastaría con las de otros estudios. No sería indispensable obtener una muestra representativa ni generalizar sus resultados. Pero al ir conociendo caso por caso, entendería las experiencias de los sujetos conquistadores atractivos y de los conquistados.

Su proceder sería inductivo: de los casos estudiados obtendría quizás el perfil que busca y el significado de conquistar.

Debemos insistir en que tanto en el proceso cuantitativo como cualitativo es posible regresar a una etapa previa. Asimismo, el planteamiento siempre es susceptible de modificarse, esto es, se encuentra en evolución.

En ambos procesos, las técnicas de recolección de los datos pueden ser múltiples. Por ejemplo, en la investigación cuantitativa: cuestionarios cerrados, registros de datos estadísticos, pruebas estandarizadas, sistemas de mediciones fisiológicas, aparatos de precisión, etc. En los estudios cualitativos: entrevistas exhaustivas, pruebas proyectivas, cuestionarios abiertos, sesiones de grupos, biografías, revisión de archivos, observación, entre otros.

Finalmente, para terminar de responder a la pregunta de este apartado, en la tabla 1.2 se comparan las etapas fundamentales de ambos procesos.

Tabla 1.2* Comparación de las etapas de investigación de los procesos cuantitativo y cualitativo

Características cuantitativas	Procesos fundamentales del proceso general de investigación	Características cualitativas
• Fundamentado en la revisión analítica de la literatura • Orientación hacia la descripción, predicción y explicación • Específico y acotado • Centrado en variables • Dirigido hacia datos medibles u observables	← Planteamiento del problema →	• Basado en la literatura y las experiencias iniciales • Orientación hacia la exploración, la descripción y el entendimiento • Emergente y abierto que va enfocándose conforme se desarrolla el proceso • Dirigido a las experiencias de los participantes

(*continúa*)

* Adaptado de Hernández-Sampieri (2014) y Creswell (2013a).

Tabla 1.2 (*continuación*)

Características cuantitativas	Procesos fundamentales del proceso general de investigación	Características cualitativas
• Direcciona el proceso • Justifica el planteamiento y la necesidad del estudio	Revisión de la literatura	• Contextualiza el proceso • Justifica el planteamiento y la necesidad del estudio
• Generalmente predeterminadas, se prueban	Hipótesis	• Generalmente emergentes
• Preestablecidos, se implementan "al pie de la letra"	Diseños	• Emergentes, se implantan de acuerdo con el contexto y circunstancias
• El tamaño depende de qué tan grande sea la población (un número representativo de casos). Se determina a partir de fórmulas y estimaciones de probabilidad	Selección de la muestra	• El tamaño depende de que comprendamos el fenómeno bajo estudio (casos suficientes). La muestra se determina de acuerdo al contexto y necesidades
• Instrumentos predeterminados • Antes de proceder al análisis se recaban todos los datos	Recolección de los datos	• Los instrumentos se van afinando • Los datos emergen paulatinamente
• Los datos encajan en categorías predeterminadas • Análisis estadístico • Descripción de tendencias, contraste de grupos o relación entre variables • Comparación de resultados con predicciones y estudios previos	Análisis de los datos	• Los datos generan categorías • Análisis temático • Descripción, análisis y desarrollo de temas • Significado profundo de los resultados
• Distribuciones de variables, coeficientes, tablas y figuras que relacionan variables, así como modelos matemáticos y estadísticos	Presentación de resultados	• Categorías, temas y patrones; tablas y figuras que asocian categorías, materiales simbólicos y modelos
• Estándar • Objetivo y sin tendencias	Reporte de resultados	• Emergente y flexible • Reflexivo y con aceptación de tendencias

¿Cuál de los dos enfoques es el mejor?

Durante décadas, las aproximaciones cuantitativa y cualitativa fueron vistas como una dicotomía (en el sentido de "rivalidad") al emprender una investigación (como "blanco y negro"), pero hoy en día la mayoría de los metodólogos las consideran como extremos en un continuo en el cual se puede situar cualquier estudio (Creswell, 2013a, Niglas, 2010). Es decir, el mundo no es nada más en blanco o negro, sino que abarca una amplísima gama de colores. Esta noción se amplía en esta obra, particularmente en el último capítulo y en varios capítulos y documentos del centro de recursos en línea.

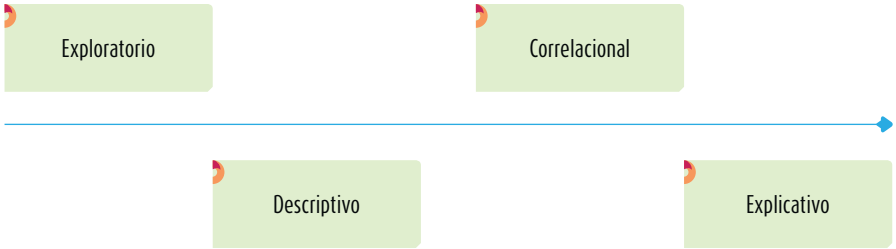
Por ello, no debemos verlos como enfoques rivales o en competencia que los investigadores pueden seguir. Desde nuestro punto de vista, ambos son muy valiosos y han servido para dar notables aportaciones al avance del conocimiento. Ninguno es intrínsecamente mejor que el otro, sólo constituyen diferentes aproximaciones al estudio de un fenómeno (como dijimos, parte de un continuo). La *investigación cuantitativa* ofrece la posibilidad de generalizar los resultados más ampliamente, otorga control sobre los fenómenos, así como un punto de vista basado en conteos y magnitudes. También, brinda una gran posibilidad de repetición y se centra en puntos específicos de tales fenómenos, además de que facilita la comparación entre estudios similares.

¿Qué alcances puede tener el proceso de investigación cuantitativa?

Si una vez hecha la revisión de la literatura hemos decidido que nuestra investigación vale la pena y debemos realizarla, el siguiente paso consiste en visualizar el alcance que tendrá.

Tal como comentamos en ediciones anteriores de este libro, **no** se deben considerar los *alcances* como "tipos" de investigación, ya que, más que ser una clasificación, constituyen un continuo de "causalidad" que puede tener un estudio, tal como se muestra en la figura 5.1.

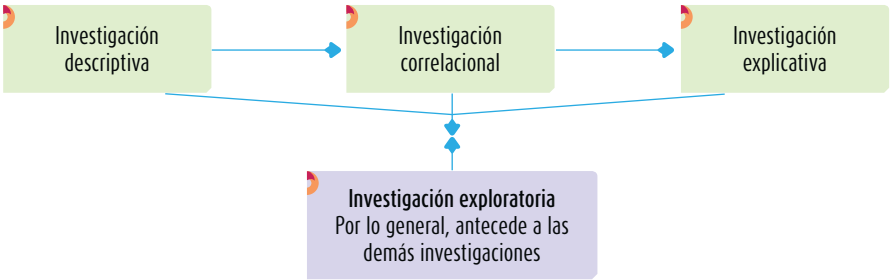
Figura 5.1 Alcances que puede tener un estudio cuantitativo.



Esta reflexión es importante, pues del alcance del estudio depende la estrategia de investigación. Así, el diseño, los procedimientos y otros componentes del proceso serán distintos en estudios con alcance exploratorio, descriptivo, correlacional o explicativo. Pero en la práctica, cualquier investigación puede incluir elementos de más de uno de estos cuatro alcances.

Los *estudios exploratorios* sirven para preparar el terreno y, por lo común, anteceden a investigaciones con alcances descriptivos, correlacionales o explicativos. Por lo general, los estudios descriptivos son la base de las investigaciones correlacionales, las cuales a su vez proporcionan información para llevar a cabo estudios explicativos que generan un sentido de entendimiento y están muy estructurados. Las investigaciones que se realizan en un campo de conocimiento específico pueden incluir diferentes alcances en las distintas etapas de su desarrollo. Es posible que una investigación se inicie como exploratoria, después puede ser descriptiva y correlacional, y terminar como explicativa (figura 5.2).

Figura 5.2 Alcances de la investigación.



Ahora bien, surge necesariamente la pregunta: ¿de qué depende que nuestro estudio se inicie como exploratorio, descriptivo, correlacional o explicativo? La respuesta no es sencilla, pero diremos que depende fundamentalmente de dos factores: el estado del conocimiento sobre el problema de investigación, mostrado por la revisión de la literatura, así como la perspectiva que se pretenda dar al estudio. Pero antes de ahondar en esta respuesta, es necesario hablar de cada uno de los alcances de la investigación.

¿En qué consisten los estudios de alcance exploratorio?

Estudios exploratorios Se emplean cuando el objetivo consiste en examinar un tema poco estudiado o novedoso.

Los **estudios exploratorios** se realizan cuando el objetivo es examinar un tema o problema de investigación poco estudiado, del cual se tienen muchas dudas o no se ha abordado antes. Es decir, cuando la revisión de la literatura reveló que tan sólo hay guías no investigadas e ideas vagamente relacionadas con el problema de estudio, o bien, si deseamos indagar sobre temas y áreas desde nuevas perspectivas.

Tal sería el caso de investigadores que pretendieran analizar fenómenos desconocidos o novedosos: una enfermedad de reciente aparición, una catástrofe ocurrida en un lugar donde nunca había sucedido algún desastre, inquietudes planteadas a partir del desciframiento del código genético humano y la clonación de seres vivos, una nueva propiedad observada en los hoyos negros del universo, el surgimiento de un medio de comunicación completamente innovador o la visión de un hecho histórico transformada por el descubrimiento de evidencia que estaba oculta.

El incremento de la esperanza de vida más allá de 100 años, la futura población que habite la Luna, el calentamiento de la atmósfera de la Tierra a niveles insospechados, cambios profundos en la concepción del matrimonio o en la ideología de una religión son hechos que generarían una gran cantidad de investigaciones exploratorias.

Los estudios exploratorios son como realizar un viaje a un sitio desconocido, del cual no hemos visto ningún documental ni leído ningún libro, sino que simplemente alguien nos hizo un breve comentario. Al llegar no sabemos qué atracciones visitar, a qué museos ir, en qué lugares se come bien, cómo es la gente; en otras palabras, ignoramos mucho del sitio. Lo primero que hacemos es explorar: preguntar sobre qué hacer y a dónde ir al taxista o al chofer del autobús que nos llevará al hotel donde nos hospedaremos; además, debemos pedir información a quien nos atienda en la recepción, al camarero, al cantinero del bar del hotel y, en fin, a cuanta persona veamos amigable. Desde luego, si ya había información del lugar y no la buscamos, perdimos dinero y mucho tiempo. De esta forma, quizá veamos un espectáculo no tan agradable y caro, al tiempo que nos perdemos de uno fascinante y más económico; por supuesto que, en el caso de la investigación científica, la inadecuada revisión de la literatura trae consecuencias más negativas que la simple frustración de gastar en algo que a fin de cuentas no nos gustó.

Valor

Los estudios exploratorios sirven para familiarizarnos con fenómenos relativamente desconocidos, obtener información sobre la posibilidad de llevar a cabo una investigación más completa respecto de un contexto particular, indagar nuevos problemas, identificar conceptos o variables promisorias, establecer prioridades para investigaciones futuras, o sugerir afirmaciones y postulados.

Esta clase de estudios son comunes en la investigación, sobre todo en situaciones en las que existe poca información. Tal fue el caso de las primeras investigaciones de Sigmund Freud, surgidas de la idea de que los problemas de histeria se relacionaban con las dificultades sexuales. Del mismo modo, los estudios pioneros del SIDA, los experimentos iniciales de Iván Pavlov sobre los reflejos condicionados y las inhibiciones, el análisis de contenido de los primeros videos musicales, las investigaciones sobre productividad de Elton Mayo en la planta Hawthorne de la compañía Western Electric, los estudios sobre terrorismo después de los atentados contra las Torres Gemelas de Nueva York en 2001, entre otros sucesos. Todos se realizaron en distintas épocas y lugares, pero con un común denominador: explorar algo poco investigado o desconocido.

Los estudios exploratorios en pocas ocasiones constituyen un fin en sí mismos. Generalmente determinan tendencias, identifican áreas, ambientes, contextos y situaciones de estudio, relaciones potenciales entre variables; o establecen el "tono" de investigaciones posteriores más elaboradas y rigurosas. Estas indagaciones se caracterizan por ser más flexibles en su método en comparación con las descriptivas, correlacionales o explicativas, y son más amplias y dispersas. Asimismo, implican un mayor "riesgo" y requieren gran paciencia, serenidad y receptividad por parte del investigador.

¿En qué consisten los estudios de alcance descriptivo?

Estudios descriptivos Busca especificar propiedades y características importantes de cualquier fenómeno que se analice. Describe tendencias de un grupo o población.

Con frecuencia, la meta del investigador consiste en describir fenómenos, situaciones, contextos y sucesos; esto es, detallar cómo son y se manifiestan. Con los **estudios descriptivos** se busca especificar las propiedades, las características y los perfiles de personas, grupos, comunidades, procesos, objetos o cualquier otro fenómeno que se someta a un análisis. Es decir, únicamente pretenden medir o recoger información de manera independiente o conjunta sobre los conceptos o las variables a las que se refieren, esto es, su objetivo no es indicar cómo se relacionan éstas. Por ejemplo, un investigador organizacional que tenga como objetivo describir varias empresas industriales de Lima, en términos de su complejidad, tecnología, tamaño, centralización y capacidad de innovación, mide estas variables y por medio de sus resultados describirá: 1) la diferenciación horizontal (subdivisión de las tareas), la vertical (número de niveles jerárquicos) y la espacial (número de centros de trabajo), así como el número de metas que han definido las empresas (complejidad); 2) qué tan automatizadas se encuentran (tecnología); 3) cuántas personas laboran en ellas (tamaño); 4) cuánta libertad en la toma de decisiones tienen los distintos niveles organizacionales y cuántos tienen acceso a la toma de decisiones (centralización de las decisiones), y 5) en qué medida llegan a modernizarse o realizar cambios en los métodos de trabajo o maquinaria (capacidad de innovación).

Sin embargo, el investigador no pretende analizar por medio de su estudio si las empresas con tecnología más automatizada son las más complejas (relacionar tecnología con complejidad) ni decirnos si la capacidad de innovación es mayor en las empresas menos centralizadas (correlacionar capacidad de innovación con centralización).

Lo mismo ocurre con el psicólogo clínico que tiene como fin describir la personalidad de un individuo. Se limitará a medir sus dimensiones (hipocondría, depresión, histeria, masculinidad-feminidad, introversión social, etc.), para posteriormente describirla. No le interesa analizar si mayor depresión se relaciona con mayor introversión social; en cambio, si pretendiera establecer relaciones entre dimensiones o asociar la personalidad con la agresividad del individuo, su estudio sería básicamente correlacional y no descriptivo.

Valor

Así como los estudios exploratorios sirven fundamentalmente para descubrir y prefigurar, los estudios descriptivos son útiles para mostrar con precisión los ángulos o dimensiones de un fenómeno, suceso, comunidad, contexto o situación.

En esta clase de estudios el investigador debe ser capaz de definir, o al menos visualizar, qué se medirá (qué conceptos, variables, componentes, etc.) y sobre qué o quiénes se recolectarán los datos (personas, grupos, comunidades, objetos, animales, hechos). Por ejemplo, si vamos a medir variables en escuelas, es necesario indicar qué tipos habremos de incluir (públicas, privadas, administradas por religiosos, laicas, de cierta orientación pedagógica, de un género u otro, mixtas, etc.). Si vamos a recolectar datos sobre materiales pétreos, debemos señalar cuáles. La descripción puede ser más o menos profunda, aunque en cualquier caso se basa en la medición de uno o más atributos del fenómeno de interés.

Ejemplo

Un censo nacional de población es un estudio descriptivo cuyo propósito es medir una serie de conceptos en un país y momento específicos: aspectos de la vivienda (tamaño en metros cuadrados, número de pisos y habitaciones, si cuenta o no con energía eléctrica y agua entubada, tipo de techo y piso, combustible utilizado, tenencia o propiedad de la vivienda, ubicación de la misma, etc.), información sobre los ocupantes (número, medios de comunicación de que disponen y edad, género, bienes, ingreso, alimentación, lugar de nacimiento, idioma o lengua, religión, nivel de estudios, ocupación de cada persona) y otras dimensiones que se juzguen relevantes para el censo. En este caso, el investigador elige una serie de conceptos a considerar, que también se denominarán variables, los mide y los resultados le sirven para describir el fenómeno de interés (la población).

Otros ejemplos de estudios descriptivos son:

1. Una investigación que determine cuál de los partidos políticos tiene más seguidores en una nación, cuántos votos ha conseguido cada uno de estos partidos en las últimas elecciones nacionales y locales, así como qué tan favorable o positiva es su imagen ante la ciudadanía.[1] Observe que no nos dice los porqués (razones).
2. Una investigación que nos indicara cuántas personas asisten a psicoterapia en una comunidad específica y a qué clase de psicoterapia acuden.

Asimismo, las características y propiedades de un conductor eléctrico, el número y tipo de divorcios anuales en una nación, el número y perfiles de pacientes que atiende un hospital, los índices de productividad generales y por área de una fábrica y la actitud hacia el aborto de un grupo de jóvenes en particular son ejemplos de información descriptiva cuyo propósito es dar un panorama (contar con una "fotografía") del fenómeno al que se hace referencia.

La revista *Rolling Stone,* dedicada a la música popular, publicó en 2012 un estudio basado en dos encuestas para determinar los 500 mejores discos de la historia (incluyendo todos los géneros de rock como pop, progresivo, alternativo, metal, punk, góspel, reggae, folk, country, etc.). La primera se efectuó en 2003, y se le pidió a un grupo de 271 músicos, productores, ejecutivos y periodistas de la industria que indicaran cuáles eran sus selecciones. La segunda se llevó a cabo en 2009 con un grupo similar de 100 expertos a fin de considerar la primera década de este siglo. A partir de los resultados de ambas se elaboró la lista de los álbumes. Asimismo, se determinó cuáles habían permanecido más semanas en el "top 40" y habían tenido más éxitos como número uno y qué artistas tenían más discos en ella. Simplemente se incluyen los álbumes de acuerdo con la frecuencia de mención y los demás datos. No se explican las razones de por qué son considerados los 500 mejores. Se trata de una investigación descriptiva.

¿En qué consisten los estudios de alcance correlacionales?

Estudio correlacional Asocian variables mediante un patrón predecible para un grupo o población.

Los **estudios correlacionales** pretenden responder a preguntas de investigación como las siguientes: ¿aumenta la autoestima de los pacientes conforme reciben una psicoterapia gestáltica? ¿A mayor variedad y autonomía en el trabajo corresponde mayor motivación intrínseca respecto de las tareas laborales? ¿Hay diferencias entre el rendimiento que otorgan las acciones de empresas de alta tecnología computacional y el rendimiento de las acciones de empresas pertenecientes a otros giros con menor grado tecnológico en la Bolsa de Comercio de Buenos Aires? ¿Los campesinos que adoptan más rápidamente una innovación son más cosmopolitas que los campesinos que la adoptan después? ¿La lejanía física entre las parejas de novios tiene una influencia negativa en la satisfacción en la relación? (Todas en un contexto específico).

Este tipo de estudios tiene como finalidad conocer la relación o grado de asociación que exista entre dos o más conceptos, categorías o variables en una muestra o contexto en particular. En ocasiones sólo se analiza la relación entre dos variables, pero con frecuencia se ubican en el estudio vínculos entre tres, cuatro o más variables.

Para evaluar el grado de asociación entre dos o más variables, en los estudios correlacionales primero se mide cada una de éstas, y después se cuantifican, analizan y establecen las vinculaciones. Tales correlaciones se sustentan en hipótesis sometidas a prueba. Por ejemplo, un investigador que desee analizar la asociación entre la motivación laboral y la productividad individual, digamos, en varias empresas industriales con más de mil trabajadores de la ciudad de Santa Fe de Bogotá, Colombia, mediría la motivación y la productividad de cada persona y después analizaría si los trabajadores con mayor motivación son o no los más productivos. Es importante recalcar que la mayoría de las veces, las mediciones de las variables que se van a correlacionar provienen de los mismos casos o participantes, pues no es lo común que se correlacionen mediciones de una variable hechas en ciertas personas, con mediciones de otra variable realizadas en personas distintas. Así, para establecer la relación entre la motivación y la productividad, no sería válido correlacionar mediciones de la motivación en trabajadores colombianos con mediciones sobre la productividad de los trabajadores peruanos.

[1] Es importante notar que la descripción del estudio puede ser más o menos general o detallada; por ejemplo, podríamos describir la imagen de cada partido político en todo el país, en cada estado, provincia o departamento; o en cada ciudad o población (y aun en los tres niveles).

Utilidad

La utilidad principal de los estudios correlacionales es saber cómo se puede comportar un concepto o una variable al conocer el comportamiento de otras variables vinculadas. Es decir, intentar predecir el valor aproximado que tendrá un grupo de individuos o casos en una variable, a partir del valor que poseen en las variables relacionadas.

Un ejemplo tal vez simple, pero que ayuda a comprender el propósito predictivo de los estudios correlacionales, sería asociar el tiempo dedicado a estudiar para un examen con la calificación obtenida. Así, en un grupo de estudiantes, se mide cuánto dedica cada uno a prepararse para el examen y también se recaban sus calificaciones (mediciones de la otra variable); luego se determina si las dos variables están relacionadas, lo cual significa que una varía cuando la otra también lo hace.

La correlación puede ser positiva o negativa. Si es positiva, significa que alumnos con valores altos en una variable tenderán también a mostrar valores elevados en la otra. Por ejemplo, quienes estudiaron más tiempo para el examen tenderían a obtener una calificación más alta. Si es negativa, significa que sujetos con valores elevados en una variable tenderán a mostrar valores bajos en la otra variable. Por ejemplo, quienes estudiaron más tiempo para el examen de estadística tenderían a obtener una calificación más baja.

Si no hay correlación entre las variables, indica que fluctúan sin seguir un patrón sistemático común; de este modo, habrá estudiantes que tengan valores altos en una de las dos variables y bajos en la otra, sujetos que tengan valores altos en una variable y altos en la otra, alumnos con valores bajos en una y bajos en la otra, y estudiantes con valores medios en las dos variables. En el ejemplo, habrá quienes dediquen mucho tiempo a estudiar para el examen y obtengan altas calificaciones, pero también quienes dediquen mucho tiempo y obtengan calificaciones bajas; otros más que dediquen poco tiempo y saquen buenas calificaciones, pero también quienes dediquen poco y les vaya mal en el examen.

Si dos variables están correlacionadas y se conoce la magnitud de la asociación, se tienen bases para predecir, con mayor o menor exactitud, el valor aproximado que tendrá un grupo de personas en una variable, al saber qué valor tienen en la otra.

Los estudios correlacionales se distinguen de los descriptivos principalmente en que, mientras que estos últimos se centran en medir con precisión las variables individuales (algunas de las cuales se pueden medir con independencia en una sola investigación), los primeros evalúan, con la mayor exactitud que sea posible, el grado de vinculación entre dos o más variables, pudiéndose incluir varios pares de evaluaciones de esta naturaleza en una sola investigación (comúnmente se incluye más de una correlación). Para comprender mejor esta diferencia, veamos un ejemplo sencillo.

Ejemplo

Supongamos que un psicoanalista tiene como pacientes a una pareja, Ana y Luis. Puede hablar de ellos de manera individual e independiente; es decir, comentar cómo es Ana (físicamente, en cuanto a su personalidad, aficiones, motivaciones, etc.) y cómo es Luis; o bien, hablar de su relación: cómo se llevan y perciben su matrimonio, cuánto tiempo pasan diariamente juntos, qué actividades comparten y otros aspectos similares. En el primer caso, la descripción es individual (si Ana y Luis fueran las variables, los comentarios del analista serían producto de un estudio descriptivo de ambos cónyuges), mientras que en el segundo, el enfoque es relacional (el interés primordial es la relación matrimonial de Ana y Luis). Desde luego, en un mismo estudio nos puede interesar tanto describir los conceptos y variables de manera individual como la relación que guardan.

Valor

La investigación correlacional tiene, en alguna medida, un valor explicativo, aunque parcial, ya que el hecho de saber que dos conceptos o variables se relacionan aporta cierta información explicativa. Por ejemplo, si la adquisición de vocabulario por parte de un grupo de niños de cierta edad (digamos entre tres y cinco años) se relaciona con la exposición a un programa de televisión educativo, ese

hecho llega a proporcionar cierto grado de explicación sobre cómo los niños adquieren algunos conceptos. Asimismo, si la similitud de valores en parejas de ciertas comunidades indígenas guatemaltecas se relaciona con la probabilidad de que contraigan matrimonio, esta información nos ayuda a explicar por qué algunas de esas parejas se casan y otras no.

Desde luego, la explicación es parcial, pues hay otros factores vinculados con la adquisición de conceptos y la decisión de casarse. Cuanto mayor sea el número de variables que se asocien en el estudio y mayor sea la fuerza de las relaciones, más completa será la explicación. En el ejemplo de la decisión de casarse, si se encuentra que, además de la similitud, también están relacionadas las variables tiempo de conocerse, vinculación de las familias de los novios, ocupación del novio, atractivo físico y tradicionalismo, el grado de explicación para la decisión de casarse será mayor. Además, si agregamos más variables que se relacionan con tal decisión, la explicación se torna más completa.

Riesgo: correlaciones espurias (falsas)

Llega a darse el caso de que dos variables estén aparentemente relacionadas, pero que en realidad no sea así. Esto se conoce en el ámbito de la investigación como correlación espuria. Supongamos que lleváramos a cabo una investigación con niños cuyas edades oscilaran entre ocho y 12 años, con el propósito de analizar qué variables se encuentran relacionadas con la inteligencia y la midiéramos por medio de alguna prueba de IQ.

Digamos también que se presenta la siguiente tendencia: a mayor estatura, mayor inteligencia; es decir que los niños físicamente más altos tendieran a obtener una calificación mayor en la prueba de inteligencia, con respecto a los niños de menor estatura. Estos resultados no tendrían sentido. No podríamos decir que la estatura se correlaciona con la inteligencia, aunque los resultados del estudio así lo indicaran.

Esto sucede por lo siguiente: la maduración está asociada con las respuestas a una prueba de inteligencia. Así, los niños de 12 años (en promedio más altos) han desarrollado mayores habilidades cognitivas para responder la prueba (comprensión, asociación, retención, etc.), que los niños de 11 años; éstos, a su vez, las han desarrollado en mayor medida que los de 10 años, y así sucesivamente hasta llegar a los niños de ocho años (en promedio los de menor estatura), quienes poseen menos habilidades que los demás para responder la prueba de inteligencia. Estamos ante una correlación espuria, cuya "explicación" no sólo es parcial sino errónea. Se requeriría de una investigación explicativa para saber cómo y por qué las variables están supuestamente relacionadas.

¿En qué consisten los estudios de alcance explicativo?

Estudios explicativos Pretenden establecer las causas de los sucesos o fenómenos que se estudian.

Los **estudios explicativos** van más allá de la descripción de conceptos o fenómenos o del establecimiento de relaciones entre conceptos; es decir, están dirigidos a responder por las causas de los eventos y fenómenos físicos o sociales. Como su nombre lo indica, su interés se centra en explicar por qué ocurre un fenómeno y en qué condiciones se manifiesta o por qué se relacionan dos o más variables.

Por ejemplo, dar a conocer las intenciones del electorado es una actividad descriptiva (indicar, según una encuesta de opinión antes de que se lleve a cabo la elección, cuántas personas "van" a votar por los candidatos contendientes constituye un estudio descriptivo) y relacionar dichas intenciones con conceptos como edad y género de los votantes o magnitud del esfuerzo propagandístico que realizan los partidos a los que pertenecen los candidatos (estudio correlacional), es diferente de señalar por qué alguien habría de votar por determinado candidato y otras personas por los demás (estudio explicativo).[2] Al hacer de nuevo una analogía con el ejemplo del psicoanalista y sus pacientes, un

[2] Como se mencionó, puede alcanzarse cierto nivel de explicación cuando: *a*) relacionamos diversas variables o conceptos y éstos se encuentran vinculados entre sí (no únicamente dos o tres, sino la mayoría de ellos), *b*) la estructura de variables presenta correlaciones considerables y, además, *c*) el investigador conoce muy bien el fenómeno de estudio. Por ahora, debido a la complejidad del tema, no se ha profundizado en algunas consideraciones sobre la explicación y la causalidad, que más adelante se expondrán.

estudio explicativo sería similar a que el médico hablara de por qué razones Ana y Luis se llevan como lo hacen (no cómo se llevan, lo cual correspondería a un nivel correlacional). Suponiendo que su matrimonio lo condujeran "bien" y la relación fuera percibida por ambos como satisfactoria, el médico explicaría por qué ocurre así. Además, nos aclararía por qué realizan ciertas actividades y pasan juntos determinado tiempo.

Ejemplo

Diferencias entre un estudio de alcance explicativo, uno descriptivo y uno correlacional

Los estudios explicativos responderían a preguntas como: ¿qué efectos tiene que los adolescentes peruanos, habitantes de zonas urbanas y de nivel socioeconómico elevado, vean videos musicales con alto contenido sexual? ¿A qué se deben estos efectos? ¿Qué variables mediatizan los efectos y de qué modo? ¿Por qué dichos adolescentes prefieren ver videos musicales con alto contenido sexual que otros programas y videos musicales? ¿Qué usos dan los adolescentes al contenido sexual de los videos musicales? ¿Qué gratificaciones derivan de exponerse a los contenidos sexuales de los videos musicales?

Un estudio descriptivo sólo respondería a preguntas como: ¿cuánto tiempo dedican esos adolescentes a ver videos musicales y especialmente videos con alto contenido sexual? ¿En qué medida les interesa ver este tipo de videos? En su jerarquía de preferencias por ciertos contenidos televisivos, ¿qué lugar ocupan los videos musicales? ¿Prefieren ver videos musicales con alto, medio, bajo o nulo contenido sexual? Por su parte, un estudio correlacional contestaría a preguntas del tipo: ¿está relacionada la exposición a videos musicales con alto contenido sexual, por parte de los mencionados adolescentes, con el control que ejercen sus padres sobre la elección de programas que hacen los jóvenes? A mayor exposición de los adolescentes a videos musicales con alto contenido sexual, ¿habrá una mayor manifestación de estrategias en las relaciones interpersonales para establecer contactos sexuales? ¿Se presentará una actitud más favorable hacia el aborto?, etcétera.

Grado de estructuración de los estudios explicativos

Las investigaciones explicativas son más estructuradas que los estudios con los demás alcances y, de hecho, implican los propósitos de éstos (exploración, descripción y correlación o asociación); además de que proporcionan un sentido de entendimiento del fenómeno a que hacen referencia.

¿Una misma investigación puede incluir diferentes alcances?

Algunas veces, una investigación puede caracterizarse como básicamente exploratoria, descriptiva, correlacional o explicativa, pero no situarse únicamente como tal. Esto es, aunque un estudio sea en esencia exploratorio, contendrá elementos descriptivos; o bien, un estudio correlacional incluirá componentes descriptivos, y lo mismo ocurre con los demás alcances.

Asimismo, debemos recordar que es posible que una investigación se inicie como exploratoria o descriptiva y después llegue a ser correlacional y aun explicativa.

Ejemplo

El hundimiento del *Titanic*[3]

Como seguramente sabrás, hace más de un siglo, el 10 de abril de 1912, el enorme buque de vapor, bautizado como *Titanic*, inició su viaje inaugural partiendo desde Southampton (Inglaterra) con destino final a Nueva York (Estados Unidos). Cuatro días más tarde, aproximadamente a las 23:40 del 14 de abril, el buque chocó con un iceberg al sur de las costas de Terranova y se hundió a las 2:20 de la mañana del 15 de abril. Murieron 1 517 personas y fue uno de los desastres marítimos más terribles en la historia (Davenport-Hines, 2012; Pontes, 2012; Titanic Inquiry Project, 2012).

[3] Tomado de Hernández-Sampieri *et al.* (2013).

capítulo 7

Concepción o elección del diseño de investigación

La gestación del diseño del estudio representa el punto donde se conectan las etapas conceptuales del proceso de investigación como el planteamiento del problema, el desarrollo de la perspectiva teórica y las hipótesis con las fases subsecuentes cuyo carácter es más operativo.

Roberto Hernández-Sampieri

Proceso de investigación cuantitativa

Paso 6 Elección o desarrollo del diseño apropiado para la investigación

- Definir cuál es el tipo de diseño más apropiado para la investigación: experimental, no experimental o múltiple.
- Precisar el diseño específico.
- Justificar el diseño elegido o desarrollado.

Objetivos de aprendizaje

Al terminar este capítulo, el alumno será capaz de:

1. Definir el significado del término "diseño de investigación", así como las implicaciones de elegir uno u otro tipo de diseño.
2. Comprender que en un estudio pueden incluirse uno o varios diseños de investigación.
3. Conocer los tipos de diseños de la investigación cuantitativa y relacionarlos con los alcances del estudio.
4. Comprender las diferencias entre la investigación experimental y la investigación no experimental.
5. Analizar los diferentes diseños experimentales y sus grados de validez.
6. Analizar los distintos diseños no experimentales y las posibilidades de investigación que ofrece cada uno.
7. Realizar experimentos y estudios no experimentales.
8. Comprender cómo el factor tiempo altera la naturaleza de un estudio.

Síntesis

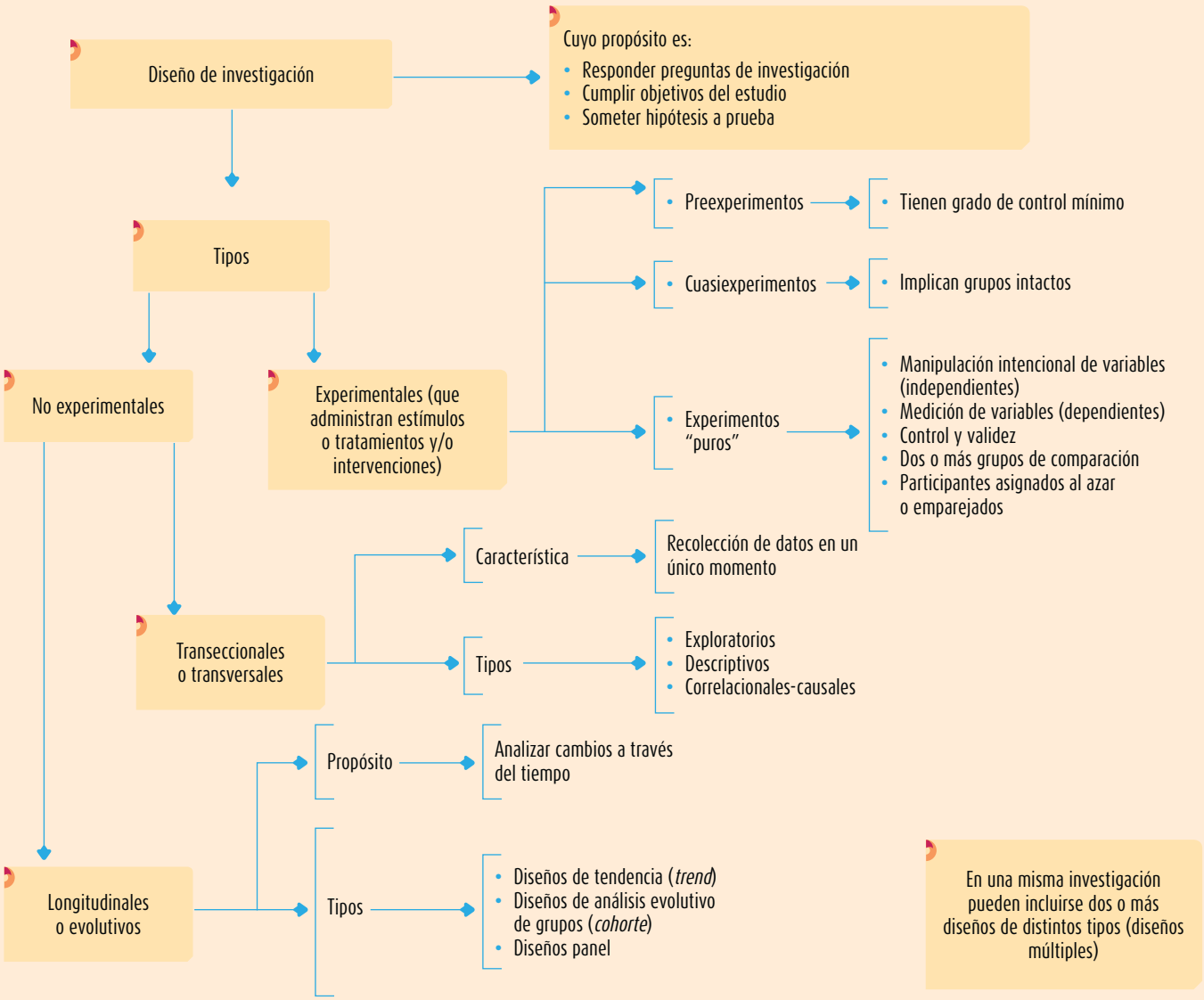
Con el propósito de responder a las preguntas de investigación planteadas y cumplir con los objetivos del estudio, el investigador debe seleccionar o desarrollar un diseño de investigación específico. Cuando se establecen y formulan hipótesis, los diseños sirven también para someterlas a prueba. Los diseños cuantitativos pueden ser experimentales o no experimentales.

En este capítulo se analizan los principales diseños experimentales y la manera de aplicarlos. Asimismo, se explica el concepto de validez experimental y cómo lograrla.

También se presenta una clasificación de diseños no experimentales, en la que se considera:

a) el factor tiempo o número de veces en que se recolectan datos.
b) el alcance del estudio.

Del mismo modo, se deja en claro que ningún tipo de diseño es intrínsecamente mejor que otro, sino que son el planteamiento del problema, los alcances de la investigación y la formulación o no de hipótesis y su tipo los que determinan qué diseño es el más adecuado para un estudio en concreto; asimismo, es posible utilizar más de un diseño.



Nota: En el centro de recursos en línea de la obra (http://www.mhhe.com/he/hmi6e) → Material complementario → Capítulos, el lector encontrará y podrá descargar el capítulo 5 adicional, "Diseños: segunda parte", que extiende los contenidos expuestos en este capítulo 7, en especial lo relativo a la técnica de asignación al azar y emparejamiento, así como a series cronológicas, diseños factoriales y cuasiexperimentos. Parte del material que estaba en ediciones anteriores en este capítulo o en el CD, se actualizó y transfirió a dicha página (***no*** se eliminó).

Centro de recursos en línea

¿Qué es un diseño de investigación?

Una vez que se precisó el planteamiento del problema, se definió el alcance inicial de la investigación y se formularon las hipótesis (o no se establecieron debido a la naturaleza del estudio), el investigador debe visualizar la manera práctica y concreta de contestar las preguntas de investigación, además de cumplir con los objetivos fijados. Esto implica seleccionar o desarrollar uno o más diseños de investigación y aplicarlos al contexto particular de su estudio. El término **diseño** se refiere al plan o estrategia concebida para obtener la información que se desea con el fin de responder al planteamiento del problema (Wentz, 2014; McLaren, 2014; Creswell, 2013a, Hernández-Sampieri *et al.*, 2013 y Kalaian, 2008).

En el enfoque cuantitativo, el investigador utiliza sus diseños para analizar la certeza de las hipótesis formuladas en un contexto en particular o para aportar evidencias respecto de los lineamientos de la investigación (si es que no se tienen hipótesis).

Sugerimos a quien se inicia en la investigación comenzar con estudios que se basen en un solo diseño y luego desarrollar indagaciones que impliquen más de uno, si es que la situación de investigación así lo requiere. Utilizar más de un diseño eleva considerablemente los costos de la investigación.

Para visualizar más claramente el asunto del diseño, recordemos una interrogante coloquial del capítulo anterior: ¿le gustaré a Paola, por qué?; y la hipótesis: "yo le resulto atractivo a Paola porque me mira frecuentemente".

Diseño Plan o estrategia que se desarrolla para obtener la información que se requiere en una investigación y responder al planteamiento.

El **diseño** constituiría el plan o la estrategia para confirmar si es o no cierto que le resulto atractivo a Paola (el plan incluiría procedimientos y actividades tendientes a encontrar la respuesta a la pregunta de investigación). En este caso podría ser: mañana buscaré a Paola después de la clase de estadística, me acercaré a ella, le diré que se ve muy guapa y la invitaré a tomar un café. Una vez que estemos en la cafetería la tomaré de la mano, y si ella no la retira, la invitaré a cenar el siguiente fin de semana; y si acepta, en el lugar donde cenemos le comentaré que me parece bonita y le preguntaré si yo le resulto atractivo. Desde luego, puedo concebir otra estrategia, tal como invitarla a bailar o ir al cine en lugar de ir a cenar; o bien, si conozco a varias amigas de Paola y yo también soy amigo de ellas, preguntarles si le gusto. En la investigación disponemos de distintas clases de diseños y debemos elegir uno o varios o desarrollar nuestra propia estrategia (por ejemplo, invitarla al cine y darle un regalo para observar cuál es su reacción al recibirlo).

Si el diseño está concebido cuidadosamente, el producto final de un estudio (sus resultados) tendrá mayores posibilidades de generar conocimiento. Y no es lo mismo seleccionar un tipo de diseño que otro: cada uno tiene sus características, como se verá más adelante. No es igual preguntarle directamente a Paola si le gusto o no, que preguntar a sus amigas; o que en lugar de interrogarla de palabra prefiera analizar su conducta no verbal (cómo me mira, qué reacciones tiene cuando la abrazo o me acerco a ella, etc.). Como tampoco será lo mismo si le pregunto delante de otras personas, que cuando estemos solos los dos. La precisión, amplitud y profundidad de la información obtenida varía en función del diseño.

¿Cómo debemos aplicar el diseño elegido o desarrollado?

En el enfoque cuantitativo, la calidad de una investigación se relaciona con el grado en que apliquemos el diseño tal como fue concebido (particularmente en el caso de los experimentos). Desde luego, en cualquier tipo de investigación el diseño se debe ajustar por contingencias o cambios en la situación (por ejemplo, en un experimento en el que no funciona el estímulo experimental, tendría que modificarse o adecuarse).

En el proceso cuantitativo, ¿de qué tipos de diseños disponemos para investigar?

En la literatura sobre la investigación cuantitativa es posible encontrar diferentes clasificaciones de los diseños. En esta obra adoptamos la siguiente clasificación:[1] investigación experimental e investigación no experimental. A su vez, la primera puede dividirse de acuerdo con las clásicas categorías de Campbell y Stanley (1966) en: preexperimentos, experimentos "puros" y cuasiexperimentos.[2] La investigación no experimental la subdividimos en diseños transversales y diseños longitudinales. En cada clasificación se comentarán los diseños específicos. De los diseños de la investigación cualitativa nos ocuparemos en la siguiente parte del libro.

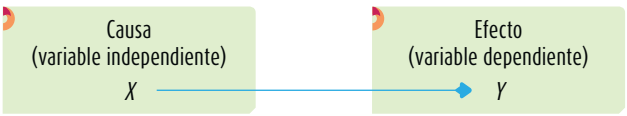
En términos generales, no consideramos que un tipo de investigación —y los consecuentes diseños— sea mejor que otro (experimental frente a no experimental). Como mencionan Kerlinger y Lee (2002), ambos son relevantes y necesarios, ya que tienen un valor propio. Cada uno posee sus características, y la decisión sobre qué clase de investigación y diseño específico hemos de seleccionar o desarrollar depende del planteamiento del problema, el alcance del estudio y las hipótesis formuladas.

Diseños experimentales

El término **experimento** tiene al menos dos acepciones, una general y otra particular. La general se refiere a "elegir o realizar una acción" y después observar las consecuencias (Babbie, 2014). Este uso del término es bastante coloquial; así, hablamos de "experimentar" cuando mezclamos sustancias químicas y vemos la reacción provocada, o cuando nos cambiamos de peinado y observamos el efecto que causa en nuestras amistades. La esencia de esta concepción de experimento es que requiere la manipulación intencional de una acción para analizar sus posibles resultados.

Una acepción particular de experimento, más armónica con un sentido científico del término, se refiere a un estudio en el que se manipulan intencionalmente una o más variables independientes (supuestas causas antecedentes), para analizar las consecuencias que la manipulación tiene sobre una o más variables dependientes (supuestos efectos consecuentes), dentro de una situación de control para el investigador (Fleiss, 2013; O'Brien, 2009 y Green, 2003). Esta definición quizá parezca compleja; sin embargo, conforme se analicen sus componentes se aclarará su sentido.

Figura 7.1 Esquema de experimento y variables.



Creswell (2013a) y Reichardt (2004) llaman a los **experimentos** estudios de intervención, porque un investigador genera una situación para tratar de explicar cómo afecta a quienes participan en ella en comparación con quienes no lo hacen. Es posible experimentar con seres humanos, seres vivos y ciertos objetos, pero siempre observando los principios éticos que se comentarán más adelante y en el capítulo 2 adicional del centro de recursos en línea.

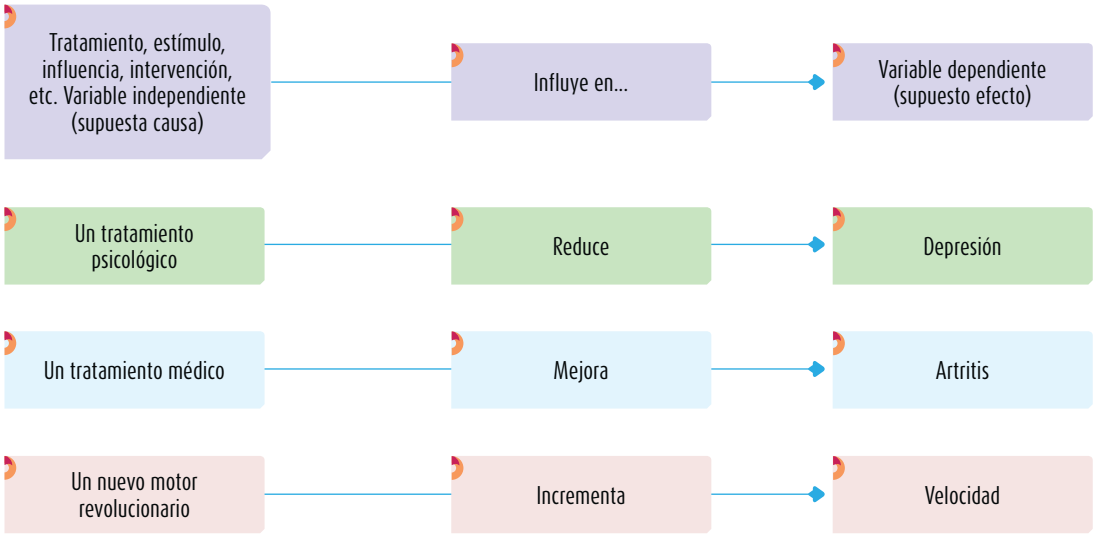
Los experimentos manipulan tratamientos, estímulos, influencias o intervenciones (denominadas variables independientes) para observar sus efectos sobre otras variables (las dependientes) en una situación de control. Veámoslo gráficamente en la figura 7.2.

[1] La tipología ha sido aceptada en ediciones anteriores por su sencillez. Otros autores han coincidido con ella como McBurney y White (2013); Creswell (2013a); Rovai, Baker y Ponton (2012), así como Kalaian (2008). Los estudios de caso se consideran como una clase diferente y de ellos se comenta más adelante, sobre todo en el capítulo 4 adicional que el lector puede descargar del centro de recursos en línea.

[2] Esta clasificación sigue siendo la más citada en textos contemporáneos; por ejemplo, Creswell (2013a), Fleiss (2013), Babbie (2014 y 2012), Silva (2008), Reichardt (2004) y Shadish, Cook y Campbell (2001).

Centro de recursos en línea

Figura 7.2 Ejemplos de la relación de variables independiente y dependiente.



Es decir, los diseños experimentales se utilizan cuando el investigador pretende establecer el posible efecto de una causa que se manipula. Pero, para establecer influencias (por ejemplo, decir que el tratamiento psicológico reduce la depresión), se deben cubrir varios requisitos que a continuación se verán.

Desde luego, hay ocasiones en que no podemos o no debemos experimentar. Por ejemplo, no podemos evaluar las consecuencias del impacto deliberadamente provocado por un meteorito sobre un planeta, pues el estímulo es imposible de manipular (¿quién puede enviar un meteorito a cierta velocidad para que choque con un planeta?). Tampoco podemos experimentar con hechos pasados, así como no debemos realizar cierto tipo de experimentos por cuestiones éticas (por ejemplo, experimentar en seres humanos con un nuevo virus para conocer su evolución). Ciertamente se han efectuado experimentos con armas bacteriológicas y bombas atómicas, castigos físicos a prisioneros, deformaciones al cuerpo humano, etc.; sin embargo, son situaciones que no deben permitirse en ninguna circunstancia.

¿Cuál es el primer requisito de un experimento?

El primer requisito es la manipulación intencional de una o más variables independientes. La variable independiente es la que se considera como supuesta causa en una relación entre variables, es la condición antecedente, y al efecto provocado por dicha causa se le denomina variable dependiente (consecuente).

Cabe destacar que el investigador puede incluir en su estudio dos o más variables independientes o dependientes. Cuando en realidad existe una relación causal entre una variable independiente y una dependiente, al variar intencionalmente la primera, la segunda también variará; por ejemplo, si la motivación es causa de la productividad, al variar la motivación deberá variar la productividad.

Experimento Situación de control en la cual se manipulan, de manera intencional, una o más variables independientes (causas) para analizar las consecuencias de tal manipulación sobre una o más variables dependientes (efectos).

Se lleva a cabo un **experimento** para analizar si una o más variables independientes afectan a una o más variables dependientes y por qué (Kirk, 2012 y Montgomery, 2012). Por ahora, simplifiquemos el problema de estudio a una variable independiente y una dependiente. En un experimento, la variable independiente resulta de interés para el investigador, ya que hipotéticamente será una de las causas que producen el efecto supuesto. Para obtener evidencia de esta supuesta relación causal, el investigador manipula la variable independiente y observa si la dependiente varía o no. Aquí, manipular es sinónimo de hacer variar o asignar distintos valores a la variable independiente.

www.elosopanda.com | jamespoetrodriguez.com

Ejemplo

Si un investigador deseara analizar el posible efecto de los contenidos televisivos antisociales en la conducta agresiva de determinados niños, podría hacer que a un grupo se le proyectara un programa de televisión con contenido antisocial y que otro viera un programa con contenido de promoción social,[3] y posteriormente observar cuál de los dos grupos muestra una conducta más agresiva.

La hipótesis de investigación nos hubiera señalado lo siguiente: "La exposición de los niños a contenidos antisociales tenderá a provocar un aumento en su conducta agresiva". De este modo, si descubre que el grupo que observó el programa antisocial muestra mayor conducta agresiva respecto del grupo que vio el programa de promoción social, y que no hay otra posible causa que hubiera afectado a los grupos de manera desigual, comprobaría su hipótesis.

El investigador manipula o hace fluctuar la variable independiente para observar el efecto en la dependiente, y lo realiza asignándole dos valores: presencia de contenidos antisociales por televisión (programa antisocial) y ausencia de contenidos antisociales por televisión (programa de promoción social). El experimentador establece la variación a propósito (no es casual): tiene el control directo sobre la manipulación y crea las condiciones para proveer la variación deseada.

En un experimento, para que una variable se considere independiente debe cumplir tres requisitos:

1. Que anteceda a la dependiente
2. Que varíe o sea manipulada
3. Que esta variación pueda controlarse

La variable dependiente se mide

La variable dependiente no se manipula, sino que se mide para ver el efecto que la manipulación de la variable independiente tiene en ella. Esto se esquematiza de la siguiente manera:

Manipulación de la variable independiente	Medición del efecto sobre la variable dependiente
X_A	Y
X_B	
.	
.	
.	

Se utiliza la letra *"X"* para simbolizar una variable independiente o tratamiento experimental. Las letras o subíndices "$_A$, $_B$…" indican distintos niveles de variación de la independiente y la letra *"Y"* se utiliza para representar una variable dependiente.

Grados de manipulación de la variable independiente

La manipulación o variación de una variable independiente puede realizarse en dos o más grados. El nivel mínimo de manipulación es de presencia o ausencia de la variable independiente. Cada nivel o grado de manipulación comprende un grupo en el experimento.

Presencia o ausencia

Este nivel o grado implica que un grupo se expone a la presencia de la variable independiente y el otro no. Posteriormente, los dos grupos se comparan para saber si el grupo expuesto a la variable independiente difiere del grupo que no fue expuesto.

Por ejemplo, a un grupo de personas con artritis se le administra el tratamiento médico y al otro grupo no se le administra. Al primero se le conoce como **grupo experimental**, y al otro, en el que está ausente la variable independiente, se le deno-

Grupo experimental Es el que recibe el tratamiento o estímulo experimental.

[3] En este momento no se explica el método para asignar a los niños a los dos grupos; lo veremos en el apartado de control y validez interna. Lo que importa ahora es que se comprenda el significado de la manipulación de la variable independiente.

Grupo de control Se le conoce también como grupo testigo.

mina **grupo de control**; pero en realidad ambos grupos participan en el experimento. Después se observa si hubo o no alguna diferencia entre los grupos en lo que respecta a la cura de la enfermedad (artritis).

A la presencia de la variable independiente se le llama "tratamiento experimental", "intervención experimental" o "estímulo experimental". Es decir, el grupo experimental recibe el tratamiento o estímulo experimental o, lo que es lo mismo, se le expone a la variable independiente; el grupo de control no recibe el tratamiento o estímulo experimental. Ahora bien, el hecho de que uno de los grupos no se exponga al tratamiento experimental no significa que su participación en el experimento sea pasiva. Por el contrario, implica que realiza las mismas actividades que el grupo experimental, excepto someterse al estímulo. En el ejemplo de la violencia televisada, si el grupo experimental va a ver un programa de televisión con contenido violento, el grupo de control podría ver el mismo programa, pero sin las escenas violentas (otra versión del programa). Si se tratara de experimentar con un medicamento, el grupo experimental consumiría el medicamento, mientras que el grupo de control consumiría un placebo (por ejemplo, una supuesta píldora que en realidad es un caramelo bajo en azúcares).

En general, en un experimento puede afirmarse lo siguiente: si en ambos grupos todo fue "igual" menos la exposición a la variable independiente, es muy razonable pensar que las diferencias entre los grupos se deban a la presencia o ausencia de tal variable.

Más de dos grados

En otras ocasiones, es posible hacer variar o manipular la variable independiente en cantidades o grados. Supongamos una vez más que queremos analizar el posible efecto del contenido antisocial por televisión sobre la conducta agresiva de ciertos niños. Podría hacerse que un grupo fuera expuesto a un programa de televisión sumamente violento (con presencia de violencia física y verbal); un segundo grupo se expusiera a un programa medianamente violento (sólo con violencia verbal), y un tercer grupo se expusiera a un programa sin violencia. En este ejemplo, se tendrían tres niveles o cantidades de la variable independiente, lo cual se representa de la siguiente manera:

X_1	(programa sumamente violento)
X_2	(programa medianamente violento)
–	(programa sin violencia, prosocial)

Manipular la variable independiente en varios niveles tiene la ventaja de que no sólo se puede determinar si la presencia de la variable independiente o tratamiento experimental tiene un efecto, sino también si distintos niveles de la variable independiente producen diferentes efectos. Es decir, si la magnitud del efecto (Y) depende de la intensidad del estímulo (X_1, X_2, X_3, etcétera).

Ahora bien, ¿cuántos niveles de variación deben ser incluidos? No hay una respuesta exacta, pues depende del planteamiento del problema y los recursos disponibles. Del mismo modo, los estudios previos y la experiencia del investigador pueden arrojar luces al respecto, ya que cada nivel implica un grupo experimental más. Por ejemplo, en el caso del tratamiento médico, dos niveles de variación pueden ser suficientes para probar su efecto, pero si tenemos que evaluar los efectos de distintas dosis de un medicamento, tendremos tantos grupos como dosis y, además, el grupo testigo o de control.

Modalidades de manipulación en lugar de grados

Hay otra forma de manipular una variable independiente que consiste en exponer a los grupos experimentales a diferentes modalidades de la variable, pero sin que esto implique cantidad. Por ejemplo, experimentar con tipos de semillas, medios para comunicar un mensaje a todos los ejecutivos de la empresa (correo electrónico o teléfono celular o memorando escrito), vacunas, estilos de argumentaciones de abogados en juicios y procedimientos de construcción o materiales.

En ocasiones, la manipulación de la variable independiente conlleva una combinación de cantidades y modalidades de ésta. Los diseñadores de automóviles experimentan con el peso del chasís

www.elosopanda.com | jamespoetrodriguez.com

(cantidad) y el material con que está construido (modalidad) para conocer su efecto en la aceleración del vehículo.

Finalmente, es necesario insistir en que cada nivel o modalidad implica, al menos, un grupo. Si hay tres niveles (grados) o modalidades, se tendrán tres grupos como mínimo.

¿Cómo se define la manera de manipular las variables independientes?

Al manipular una variable independiente es necesario especificar qué se va a entender por esa variable en el experimento (definición operacional experimental). Es decir, trasladar el concepto teórico a un estímulo experimental. Por ejemplo, si la variable independiente a manipular es la exposición a la violencia televisada (en adultos), el investigador debe pensar cómo va a transformar ese concepto en una serie de operaciones experimentales. En este caso podría ser: la violencia televisada será operacionalizada (transportada a la realidad) mediante la exposición a un programa donde haya riñas y golpes, insultos, agresiones, uso de armas de fuego, crímenes e intentos de crímenes, ataques sexuales, intimidación, persecuciones, etc. Entonces se selecciona un programa en el que se muestren tales conductas (por ejemplo, *CSI*: *Investigación de la escena del crimen, Hannibal o La ley y el orden*: *Unidad de víctimas especiales*, o una telenovela o serie producida en Iberoamérica en que se presenten dichos comportamientos). Así, el concepto abstracto se transforma en un referente real.

Veamos cómo un concepto teórico (grado de información sobre la deficiencia mental) en la práctica se tradujo a dos niveles de manipulación experimental.

Ejemplo

Naves y Poplawsky (1984) diseñaron un experimento para poner a prueba la siguiente hipótesis: "a mayor grado de información sobre la deficiencia mental que posea el sujeto común se mostrará menor evitación en la interacción con el deficiente mental".[4]

La variable independiente fue "el grado de información sobre la deficiencia mental" (o mejor dicho, capacidad mental distinta); y la dependiente, "la conducta de evitación en interacciones con personas cuyas capacidades mentales son diferentes". La primera se manipuló mediante dos niveles de información: 1) información cultural, y 2) información sociopsicológica acerca de esta capacidad mental. Por tanto, hubo dos grupos: uno con información cultural y otro con información sociopsicológica. El primer grupo no recibió ninguna información sobre la deficiencia mental o la capacidad mental distinta, ya que se supuso "que todo individuo, por pertenecer a cierta cultura, maneja este tipo de información, y está conformada por nociones generales y normalmente estereotipadas sobre la deficiencia mental; de ello se desprende que si un sujeto basa sus predicciones sobre la conducta del otro en el nivel cultural, obtendrá mínima precisión y pocas probabilidades de controlar el evento comunicativo" (Naves y Poplawsky, 1984, p. 119).

El segundo grupo acudió a un centro de entrenamiento para personas cuyas capacidades mentales son diferentes, quienes les proporcionaron información sociopsicológica (algunas contaron sus problemas en el trabajo y sus relaciones con superiores y compañeros, también se trataron temas como el amor y la amistad). Este grupo pudo observar lo que es la "capacidad mental distinta", cómo se trata clínicamente y los efectos en la vida cotidiana de quien la posee, además de recibir información sociopsicológica al respecto.

Después, todos los participantes fueron expuestos a una interacción sorpresiva con un supuesto individuo con capacidad mental distinta (que en realidad era un actor entrenado para comportarse como "deficiente mental" y con conocimientos sobre la materia).[5] La situación experimental estuvo bajo riguroso control y se filmaron las interacciones para medir el grado de evitación hacia el sujeto con capacidad mental diferente, a través de cuatro dimensiones: *a*) distancia física, *b*) movimientos corporales que denotaban tensión, *c*) conducta visual y *d*) conducta verbal. Se comprobó la hipótesis, pues el grupo con información cultural mostró una mayor conducta de evitación que el grupo con información sociopsicológica.

[4] En el ejemplo a veces se emplean los términos "deficiencia mental" y "deficiente mental", debido a que son los que utilizaron Esther Naves y Silvia Poplawsky. Tal vez serían más adecuados los términos: "capacidad mental diferente" y "persona con tal capacidad". De antemano, una disculpa si alguien se siente ofendido por estos vocablos.

[5] Las actuaciones fueron ensayadas una y otra vez ante un grupo de cuatro expertos sobre la deficiencia mental, hasta que el grupo validó unánimemente el desempeño del actor.

Dificultades para definir cómo se manipularán las variables independientes

En ocasiones no resulta tan difícil trasladar el concepto teórico (variable independiente) a operaciones prácticas de manipulación (tratamientos o estímulos experimentales). Manipular la paga (cantidades de dinero otorgadas), la realimentación, el reforzamiento y la administración de un medicamento no es demasiado complejo. Sin embargo, a veces resulta verdaderamente complicado representar el concepto teórico en la realidad, sobre todo con variables internas, variables que pueden tener diversos significados o variables que sean difíciles de alterar. La socialización, la cohesión, la conformidad, el poder, la motivación individual y la agresión son conceptos que requieren un enorme esfuerzo por parte del investigador para operacionalizarse.

Guía para sortear dificultades

Para definir cómo se va a manipular una variable es necesario:

1. *Consultar experimentos antecedentes* para ver si fue exitosa la forma de manipular la variable independiente. Al respecto, resulta imprescindible analizar si la manipulación de esos estudios puede aplicarse al contexto específico del nuestro, o cómo se extrapolaría a nuestra situación experimental.
2. *Evaluar la manipulación* antes de que se conduzca el experimento. Hay varias preguntas que el experimentador debe hacerse para evaluar su manipulación antes de llevarla a cabo: ¿las operaciones experimentales representan la variable conceptual que se tiene en mente? ¿Los diferentes niveles de variación de la variable independiente harán que los sujetos se comporten de diferente forma? (Christensen, 2006). ¿Qué otras maneras hay para manipular la variable? ¿Ésta es la mejor? Si el concepto teórico no se traslada adecuadamente a la realidad, lo que sucederá es que al final realizaremos otro experimento muy distinto del que pretendemos. Si deseáramos averiguar el efecto de la ansiedad sobre la memorización de conceptos y si nuestra manipulación es errónea (en lugar de provocar ansiedad, generase inconformidad), los resultados del experimento tal vez nos ayudarían a explicar la relación entre la inconformidad y la memorización de conceptos; pero de ninguna manera servirán para analizar el efecto de la ansiedad en la memorización.

   También, si la presencia de la variable independiente en los grupos experimentales es "débil", probablemente no se encontrarán efectos, pero no porque no pueda haberlos. Si pretendemos manipular la violencia televisada y nuestro programa no es en realidad violento (incluye uno que otro insulto y algunas sugerencias de violencia física) y no encontramos un efecto, no podemos afirmar o negar que haya un efecto, porque la manipulación fue débil.
3. *Incluir verificaciones para la manipulación*. Cuando se experimenta con personas hay varias formas de verificar si realmente funcionó la manipulación. La primera consiste en entrevistar a los participantes. Supongamos que, por medio de la manipulación, pretendemos generar que un grupo esté muy motivado hacia una tarea, y el otro no. Después del experimento entrevistaríamos a los individuos para ver si el grupo que debía estar muy motivado en realidad lo estuvo y si el grupo que no debía estar motivado no lo estuvo. Una segunda forma es incluir mediciones relativas a la manipulación durante el experimento. Por ejemplo, aplicar una escala de motivación a ambos grupos cuando supuestamente unos deben estar motivados y otros no.

¿Cuál es el segundo requisito de un experimento?

El segundo requisito consiste en medir el efecto que la variable independiente tiene en la variable dependiente. Esto es igualmente importante y como en la variable dependiente se observa el efecto, la medición debe ser adecuada, válida y confiable.

Imaginemos que conducimos un experimento para evaluar el efecto de un nuevo tipo de enseñanza en la comprensión de conceptos políticos por parte de ciertos niños, y en lugar de medir comprensión medimos la memorización; por más correcta que resulte la manipulación de la variable

www.elosopanda.com | jamespoetrodriguez.com

independiente, el experimento resultaría un fracaso porque la medición de la dependiente no es válida. O pensemos que tenemos dos grupos a comparar con mediciones distintas, y si encontramos diferencias ya no sabremos si se debieron a la manipulación de la independiente o a que se aplicaron exámenes de comprensión diferentes. Los requisitos para medir correctamente una variable se comentan en el capítulo 9: "Recolección de los datos cuantitativos".

¿Cuántas variables independientes y dependientes deben incluirse en un experimento?

No hay reglas para ello; depende de cómo se haya planteado el problema de investigación y las limitaciones existentes. Si un investigador se encuentra interesado en contrastar efectos de apelaciones emotivas frente a racionales en comerciales televisivos sobre la predisposición de compra de un producto, y sólo le interesa este problema, tendrá una variable independiente única y una sola dependiente. Pero si también pretende analizar el efecto de utilizar comerciales en blanco y negro frente a los que son a color, agregaría esta variable independiente y la manipularía. Tendría dos variables independientes (apelación y colorido) y una dependiente (predisposición de compra), son cuatro grupos (sin contar el de control):

a) Grupo expuesto a apelación emotiva y comercial en blanco y negro
b) Grupo expuesto a apelación emotiva y comercial en color
c) Grupo expuesto a apelación racional y comercial en blanco y negro
d) Grupo expuesto a apelación racional y comercial en color

O también se podría agregar una tercera variable independiente: duración de los comerciales, y una cuarta: realidad de los modelos del comercial (personas vivas en contraposición a dibujos animados) y así sucesivamente. Claro está que conforme se aumenta el número de variables independientes se incrementan las manipulaciones que deben hacerse y el número de grupos requerido. Entonces, entraría en juego el segundo factor: limitantes, tal vez no pueda reclutar las suficientes personas para varios grupos o contar con el presupuesto para producir tal variedad de comerciales.

Por otro lado, en cada caso podría optar por medir más de una variable dependiente y evaluar múltiples efectos de las independientes (en distintas variables). Por ejemplo, además de la predisposición de compra, medir la recordación del comercial y la evaluación estética de éste. Resulta obvio que al aumentar las variables dependientes, no tienen que incrementarse los grupos, porque estas variables no se manipulan. Lo que aumenta es el tamaño de la medición (cuestionarios con más preguntas, mayor número de observaciones, entrevistas más largas, etc.) porque hay más variables que medir.

¿Cuál es el tercer requisito de un experimento?

El tercer requisito es el control o la **validez interna** de la situación experimental. El término "control" tiene diversas connotaciones. Sin embargo, su acepción más común es que, si en el experimento se observa que una o más variables independientes hacen variar a las dependientes, la variación de estas últimas se debe a la manipulación de las primeras y no a otros factores o causas; y si se observa que una o más independientes no tienen un efecto sobre las dependientes, se puede estar seguro de ello. Es decir, saber qué está ocurriendo realmente con la relación entre las variables independientes y las dependientes. Esto podría ilustrarse de la siguiente manera:

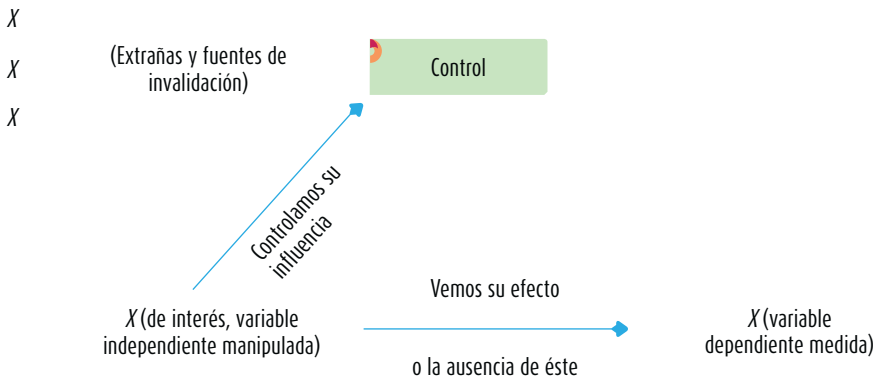
Validez interna Grado de confianza que se tiene de que los resultados del experimento se interpreten adecuadamente y sean válidos (se logra cuando hay control).

Figura 7.3 Experimentos con control e intento de experimento.

Cuando hay **control** es posible determinar la relación causal; cuando no se logra el control, no se puede conocer dicha relación (no se sabe qué está detrás del "cuadro en color", quizá sería, por ejemplo: "X → Y", o "X | Y"; es decir, que hay correlación o que no existe ninguna relación).

Así, lograr *control* en un experimento implica contener la influencia de otras variables extrañas en las variables dependientes, para conocer en realidad si las variables independientes que nos interesan tienen o no efecto en las dependientes. Lo anterior se esquematizaría así:

Figura 7.4 Experimentos con control de las variables extrañas.



Es decir, "purificamos" la relación de *X* (independiente) con *Y* (dependiente) de otras posibles fuentes que afecten a *Y*, y que "contaminen" el experimento. Aislamos las relaciones que nos interesan. Si deseamos analizar el efecto que pueda tener un comercial sobre la predisposición de compra hacia el producto que se anuncia, sabemos que quizás existan otras razones o causas por las cuales las personas piensen en comprar el producto (calidad, precio, cualidades, prestigio de la marca, etc.). Entonces, en el experimento se deberá controlar la posible influencia de estas otras causas, para que sepamos si el comercial tiene o no algún efecto. De lo contrario, si se observa que la predisposición de compra es elevada y no hay control, no sabremos si el comercial es la causa o lo son los demás factores.

Lo mismo ocurre con un método de enseñanza, cuando por medio de un experimento se desea evaluar su influencia en el aprendizaje. Si no hay control, no sabremos si un buen aprendizaje se debió al método, a que los participantes eran sumamente inteligentes, a que tenían conocimientos aceptables de los contenidos o a cualquier otro motivo. Si no hay aprendizaje no sabremos si se debe a que los sujetos estaban muy desmotivados con los contenidos, a que eran poco inteligentes o a cualquier otra causa. Es decir, buscamos descartar otras posibles explicaciones para evaluar si la nuestra es o no la correcta (variables independientes de interés, estímulos o tratamientos experimentales que tienen el efecto que nos interesa comprobar). Tales explicaciones rivales son las fuentes de invalidación interna (que pueden invalidar el experimento).

Fuentes de invalidación interna

Existen diversos factores que tal vez nos confundan y sean motivo de que ya no sepamos si la presencia de una variable independiente o un tratamiento experimental surte o no un verdadero efecto. Se trata de explicaciones rivales de la explicación de que las variables independientes afectan a las dependientes (The SAGE Glossary of the Social and Behavioral Sciences, 2009a). Campbell y Stanley (1966) definieron estas explicaciones rivales, las cuales han sido ampliadas por Campbell (1975), Christensen (2006), Hernández-Sampieri *et al.* (2013) y Babbie (2014). Se les conoce como fuentes de invalidación interna porque precisamente atentan contra la validez interna de un experimento. Ésta se refiere a cuánta confianza tenemos en que sea posible interpretar los resultados del experimento y éstos sean válidos. La *validez interna* se relaciona con la calidad del experimento y se logra cuando hay control, cuando los grupos difieren entre sí solamente en la exposición a la variable independiente (ausencia o presencia o en grados o modalidades), cuando las mediciones de la variable

dependiente son confiables y válidas y cuando el análisis es el adecuado para el tipo de datos que estamos manejando. El control en un experimento se alcanza eliminando esas explicaciones rivales o fuentes de invalidación interna. En la tabla 7.1 se mencionan algunas fuentes. Una explicación más amplia, así como ejemplos y otras fuentes posibles, las podrá encontrar el lector en el capítulo 5 adicional, "Diseños: segunda parte", que se puede descargar del centro de recursos en línea.

Centro de recursos en línea

Tabla 7.1 Principales fuentes de invalidación interna[6]

Fuente o amenaza a la validez interna	Descripción de la amenaza	En respuesta, el investigador debe:
Historia	Eventos o acontecimientos externos que ocurran durante el experimento e influyan solamente a algunos de los participantes.	Asegurarse de que los participantes de los grupos experimentales y de control experimenten los mismos eventos.
Maduración	Los participantes pueden cambiar o madurar durante el experimento y esto afectar los resultados.	Seleccionar participantes para los grupos que maduren o cambien de manera similar durante el experimento.
Inestabilidad del instrumento de medición	Poca o nula confiabilidad del instrumento.	Elaborar un instrumento estable y confiable.
Inestabilidad del ambiente experimental	Las condiciones del ambiente o entorno del experimento no sean iguales para todos los grupos participantes.	Lograr que las condiciones ambientales sean las mismas para todos los grupos.
Administración de pruebas	Que la aplicación de una prueba o instrumento de medición antes del experimento influya las respuestas de los individuos cuando se vuelve a administrar la prueba después del experimento (por ejemplo, recuerden sus respuestas).	Tener pruebas equivalentes y confiables, pero que no sean las mismas y que los grupos que se comparen sean equiparables.
Instrumentación	Que las pruebas o instrumentos aplicados a los distintos grupos que participan en el experimento no sean equivalentes.	Administrar la misma prueba o instrumento a todos los individuos o grupos participantes.
Regresión	Seleccionar participantes que tengan puntuaciones extremas en la variable medida (casos extremos) y que no se mida su valoración real.	Elegir participantes que no tengan puntuaciones extremas o pasen por un momento anormal.
Selección	Que los grupos del experimento no sean equivalentes.	Lograr que los grupos sean equivalentes.
Mortalidad	Que los participantes abandonen el experimento.	Reclutar suficientes participantes para todos los grupos.
Difusión de tratamientos	Que los participantes de distintos grupos se comuniquen entre sí y esto afecte los resultados.	Durante el experimento mantener a los grupos tan separados entre sí como sea posible.
Compensación	Que los participantes del grupo de control perciban que no reciben nada y eso los desmoralice y afecte los resultados.	Proveer de beneficios a todos los grupos participantes.
Conducta del experimentador	Que el comportamiento del experimentador afecte los resultados.	Actuar igual con todos los grupos y ser "objetivo".

¿Cómo se logran el control y la validez interna?

El control en un experimento logra la validez interna y se alcanza mediante:

1. Varios grupos de comparación (dos como mínimo).
2. Equivalencia de los grupos en todo, excepto en la manipulación de la o las variables independientes.

[6] Basada en Hernández-Sampieri *et al.* (2013) y Mertens (2010), pero principalmente en Creswell (2013a).

Es necesario que en un experimento se tengan, por lo menos, dos grupos que comparar. Si nada más se tiene un grupo no es posible saber con certeza si influyeron las fuentes de invalidación interna u otras causas ajenas a la variable independiente manipulada. Imaginemos un experimento en el cual queremos ver si un fertilizante hace crecer más rápidamente cierta especie de plantas. Debemos tener plantas a las cuales se les administre el fertilizante (grupo experimental) y plantas a las que no (grupo de control). Siempre debe existir un punto de comparación.

Pero no basta con dos o más grupos, sino que éstos deben ser similares en todo, menos en la manipulación de la o las variables independientes. El control implica que todo permanece constante, salvo tal manipulación o intervención. Si entre los grupos que conforman el experimento todo es similar o equivalente, excepto la manipulación de la variable independiente, las diferencias entre los grupos pueden atribuirse a ella y no a otros factores.

Veamos un ejemplo educativo, supongamos que deseamos analizar si una serie de programas educativos de televisión para niños genera mayor aprendizaje de los preceptos básicos de investigación en comparación con un método educativo tradicional. Un grupo recibe la enseñanza a través de los programas, otro grupo la obtiene por medio de instrucción oral tradicional y un tercer grupo dedica ese mismo tiempo a actividades recreativas en el salón de clases. Asumamos que los niños que aprendieron mediante los programas obtienen las mejores calificaciones en una prueba de conocimientos relativa a los contenidos enseñados, los que recibieron el método tradicional obtienen calificaciones mucho más bajas, y los que jugaron obtienen puntuaciones de cero o cercanas a este valor. En forma aparente, los programas son un mejor vehículo de enseñanza que la instrucción oral. Pero si los grupos no son equivalentes, entonces no podemos confiar en que las diferencias se deban en realidad a la manipulación de la variable independiente (programas televisivos-instrucción oral) y no a otros factores, o a la combinación de ambos. Por ejemplo, a los niños más inteligentes, estudiosos y con mayor empeño se les asignó al grupo que fue instruido por televisión, o la instructora del método tradicional no tenía buen desempeño, o los niños expuestos a este último método recibieron mayor carga de trabajo y tenían exámenes los días en que se desarrolló el experimento, etc. ¿Cuánto se debió al método y cuánto a otros factores? Para el investigador la respuesta a esta pregunta se convierte en un enigma: no hay control.

Los grupos deben ser equivalentes al inicio del experimento y durante su desarrollo, salvo en lo que respecta a la variable independiente. Asimismo, los instrumentos de medición deben ser iguales y aplicados de la misma manera.

¿Cómo se logra la equivalencia inicial?: asignación al azar

Existe un método muy difundido para alcanzar esta equivalencia: la **asignación aleatoria o al azar** de los participantes a los grupos del experimento.[7] La **asignación al azar** asegura probabilísticamente que dos o más grupos son equivalentes entre sí (Kirk, 2012; Knapp, 2008; Pettygrove, 2007 y Peng, 2003). Es una técnica de control que tiene como propósito dar al investigador la seguridad de que variables extrañas, conocidas o desconocidas, no afectarán de manera sistemática los resultados del estudio (Christensen, 2006).

La asignación al azar puede llevarse a cabo empleando trozos de papel. Se escribe el nombre de cada caso o participante (o alguna clave que lo identifique) en los papeles, los cuales se juntan en algún recipiente, se revuelven y se van sacando —sin observarlos— para formar los grupos. Por ejemplo, si se tienen dos grupos, las personas con turno non en su papel irían al primer grupo; y las personas con par, al segundo grupo. O bien, si hubiera 80 personas, los primeros 40 papelitos que se saquen irían a un grupo y los restantes 40 al otro.

Asignación aleatoria o al azar Es una técnica de control muy difundida para asegurar la equivalencia inicial al ser asignados aleatoriamente los casos o sujetos a los grupos del experimento.

También, cuando se tienen dos grupos, la **asignación aleatoria** puede llevarse a cabo utilizando una moneda no cargada. Se lista a los participantes y se designa qué lado de la moneda va a significar el grupo uno y qué lado el grupo dos. Con cada sujeto se lanza la moneda y, dependiendo del resultado, se asigna a uno u otro grupo.

[7] El que los participantes sean asignados al azar significa que no hay un motivo sistemático por el cual fueron elegidos para ser parte de un grupo o de otro, sino que la casualidad es lo que define a qué grupo son asignados.

www.elosopanda.com | jamespoetrodriguez.com

Tal procedimiento está limitado sólo a dos grupos, porque las monedas tienen dos caras, aunque podrían utilizarse dados o cubos, por ejemplo.

Una tercera forma de asignar los participantes o casos a los grupos es mediante el programa STATS®, seleccionando el subprograma "Generador de números aleatorios". Previamente se numeran todos los casos (supongamos que se tiene un experimento con dos grupos y 100 personas en total, consecuentemente se numera a los sujetos del 1 al 100). El programa pregunta en la ventana: ¿Cuántos números aleatorios desea generar? Entonces se escribe el número relativo al total de los participantes en el experimento, así, debe teclearse "100". Inmediatamente se elige la opción: Valores mínimo y máximo para los números aleatorios, en el mínimo se introduce un "1" (siempre será "1") y en el máximo un "100" (o el número total de participantes). Posteriormente se elige como medio de sorteo "random" (al azar) y se hace clic en Calcular y el programa generará 100 números de manera aleatoria, así, se pueden asignar los primeros 50 a un grupo y los últimos 50 al otro grupo, o bien, el primer número al grupo 1, el segundo al grupo 2, el tercero al grupo 1 y así sucesivamente (dado que la generación de los números es completamente aleatoria, en ocasiones el programa duplica o triplica algunos números, entonces debemos saltarnos uno o dos de los números repetidos y seguir asignando sujetos —números— a los grupos; y al terminar se vuelve a repetir el proceso y continuamos asignando a los grupos los números que no habían "salido" antes, hasta situar a los 100 sujetos en los dos grupos (si fueran cuatro grupos, los primeros 25 se asignan al grupo 1, los segundos 25 al grupo 2, los siguientes 25 al grupo 3 y los últimos 25 al grupo 4).

La **asignación al azar** produce control, pues las variables que deben ser controladas (variables extrañas y fuentes de invalidación interna) se distribuyen aproximadamente de la misma manera en los grupos del experimento. Y puesto que la distribución es bastante similar en todos los grupos, la influencia de otras variables que no sean la o las independientes se mantiene constante, porque aquellas no pueden ejercer ninguna influencia diferencial en las variables dependientes (Christensen, 2006).

La asignación aleatoria funciona mejor cuanto mayor sea el número de casos con que se cuenta para el experimento, es decir, cuanto mayor sea el tamaño de los grupos. Se recomienda que para cada grupo se tengan por lo menos 15 personas.[8]

Otra técnica para lograr la equivalencia inicial: el emparejamiento

Un método alternativo para intentar hacer inicialmente equivalentes a los grupos es el **emparejamiento** o la **técnica de apareo** (en inglés, *matching*). El proceso consiste en igualar a los grupos en relación con alguna variable específica que puede influir de modo decisivo en la o las variables dependientes.

Técnica de apareo o emparejamiento
Consiste en igualar a los grupos en relación con alguna variable específica, que puede influir de modo decisivo en la variable dependiente.

El primer paso es elegir la variable concreta de acuerdo con algún criterio teórico. Es obvio que esta variable debe estar muy relacionada con las variables dependientes. Si se pretendiera analizar el efecto que causa utilizar distintos tipos de materiales suplementarios de instrucción sobre el desempeño en la lectura, el apareamiento podría basarse en la variable "agudeza visual". Experimentos sobre métodos de enseñanza emparejarían a los grupos en "conocimientos previos", "aprovechamiento anterior en una asignatura relacionada con los contenidos a enseñar" o "inteligencia". Experimentos sobre las actitudes hacia productos o hábitos de compra pueden utilizar la variable "ingreso" para empatar los grupos. En cada caso en particular, debe pensarse cuál es la variable cuya influencia sobre los resultados del experimento resulta más necesario controlar y buscar el apareamiento de los grupos en esa variable.

El segundo paso consiste en obtener una medición de la variable elegida para emparejar a los grupos. Esta medición puede existir o efectuarse antes del experimento. Vamos a suponer que nuestro experimento fuera sobre métodos de enseñanza, el emparejamiento llegaría a hacerse sobre la base de la inteligencia. Si fueran adolescentes, se obtendrían registros de inteligencia de ellos o se les aplicaría una prueba de inteligencia.

[8] Este criterio se basa en los requisitos de algunos análisis estadísticos.

Asignación al azar Es el mejor método para hacer equivalentes los grupos (más preciso y confiable). El emparejamiento no la sustituye por completo.

El tercer paso es ordenar a los participantes en la variable sobre la cual se va a efectuar el emparejamiento (de las puntuaciones más altas a las más bajas).

El cuarto paso consiste en formar parejas, tercias, cuartetos, etc., de participantes según la variable de apareamiento (son individuos que tienen la misma puntuación en la variable o una calificación similar) e ir asignando a cada integrante de cada pareja, tercia o similar a los grupos del experimento, buscando un equilibrio entre éstos. También podría intentarse empatar a los grupos en dos variables, pero ambas deben estar sumamente relacionadas, porque de lo contrario resultaría muy difícil el emparejamiento. Conforme más variables se utilizan para aparear grupos, el procedimiento es más complejo. En el capítulo 5 adicional: "Diseños: segunda parte" se ejemplifica el procedimiento.

Centro de recursos en línea

Una tipología sobre los diseños experimentales

A continuación se presentan los diseños experimentales más citados en la literatura. Para ello nos basaremos, como ya se señaló, en la tipología de Campbell y Stanley (1966), quienes dividen los diseños experimentales en tres clases: *a*) preexperimentos, *b*) experimentos "puros"[9] y *c*) cuasiexperimentos. Se utilizará la simbología que generalmente se emplea en los textos sobre experimentos.

Simbología de los diseños experimentales

Aquí se describe una simbología básica para diseños experimentales:

- *R* Asignación al azar o aleatoria. Cuando aparece quiere decir que los sujetos han sido asignados a un grupo de manera aleatoria (proviene del inglés *randomization,* "aleatorización").
- *G* Grupo de sujetos o casos (G_1, grupo 1; G_2, grupo 2; etcétera).
- *X* Tratamiento, estímulo o condición experimental (presencia de algún nivel o modalidad de la variable independiente).
- 0 Una medición de los sujetos de un grupo (prueba, cuestionario, observación, etc.). Si aparece antes del estímulo o tratamiento, se trata de una preprueba (previa al tratamiento). Si aparece después del estímulo se trata de una posprueba (posterior al tratamiento).
- — Ausencia de estímulo (nivel "cero" en la variable independiente). Indica que se trata de un grupo de control o testigo.

Asimismo, cabe mencionar que la secuencia horizontal indica tiempos distintos (de izquierda a derecha) y cuando en dos grupos aparecen dos símbolos alineados verticalmente, esto indica que tienen lugar en el mismo momento del experimento. Veamos de manera gráfica estas dos observaciones.

Figura 7.5 Simbología de los diseños experimentales.

RG_1	0	*X*	0
Primero, se asigna a los participantes al azar al grupo 1	Segundo, se aplica una medición previa	Tercero, se administra el estímulo	Cuarto, se aplica una medición posterior

RG_1	*X*	0
RG_2	–	0

Ambos símbolos están alineados verticalmente, lo cual significa que tienen lugar en el mismo momento.

[9] Preferimos utilizar el término "experimentos puros" más que "verdaderos" (que es el término original y así se ha traducido en diversas obras), porque crea confusión entre los estudiantes.

Preexperimentos

Los **preexperimentos** se llaman así porque su grado de control es mínimo.

1. Estudio de caso con una sola medición

Este diseño podría diagramarse de la siguiente manera:

$$G \qquad X \qquad 0$$

Consiste en administrar un estímulo o tratamiento a un grupo y después aplicar una medición de una o más variables para observar cuál es el nivel del grupo en éstas.

Este diseño no cumple con los requisitos de un experimento "puro". No hay manipulación de la variable independiente (niveles) o grupos de contraste (ni siquiera el mínimo de presencia o ausencia). Tampoco hay una referencia previa de cuál era el nivel que tenía el grupo en la o las variables dependientes antes del estímulo. No es posible establecer causalidad con certeza ni se controlan las fuentes de invalidación interna.

2. Diseño de preprueba/posprueba con un solo grupo

Este segundo diseño se diagramaría así:

$$G \qquad 0_1 \qquad X \qquad 0_2$$

A un grupo se le aplica una prueba previa al estímulo o tratamiento experimental, después se le administra el tratamiento y finalmente se le aplica una prueba posterior al estímulo.

Este diseño ofrece una ventaja sobre el anterior: existe un punto de referencia inicial para ver qué nivel tenía el grupo en las variables dependientes antes del estímulo; es decir, hay un seguimiento del grupo. Sin embargo, el diseño no resulta conveniente para fines de establecer causalidad: no hay manipulación ni grupo de comparación y es posible que actúen varias fuentes de invalidación interna, por ejemplo, la historia. Entre 0_1 y 0_2 podrían ocurrir otros acontecimientos capaces de generar cambios, además del tratamiento experimental, y cuanto más largo sea el lapso entre ambas mediciones, mayor será también la posibilidad de que actúen tales fuentes.

Por otro lado, se corre el riesgo de elegir a un grupo atípico o que en el momento del experimento no se encuentre en su estado normal.

En ocasiones este diseño se utiliza con un solo individuo (estudio de caso experimental). Sobre tal diseño se abunda en el capítulo cuatro adicional: "Estudios de caso", el cual se puede descargar del centro de recursos en línea.

Centro de recursos en línea

Los dos diseños preexperimentales no son adecuados para el establecimiento de relaciones causales porque se muestran vulnerables en cuanto a la posibilidad de control y validez interna. Algunos autores consideran que deben usarse sólo como ensayos de otros experimentos con mayor control.

En ciertas ocasiones los **diseños preexperimentales** sirven como estudios exploratorios, pero sus resultados deben observarse con precaución.

Diseño preexperimental Diseño de un solo grupo cuyo grado de control es mínimo. Generalmente es útil como un primer acercamiento al problema de investigación en la realidad.

Experimentos "puros"

Los experimentos "puros" son aquellos que reúnen los dos requisitos para lograr el control y la validez interna:

1. Grupos de comparación (manipulación de la variable independiente).
2. Equivalencia de los grupos.

Estos diseños llegan a incluir una o más variables independientes y una o más dependientes. Asimismo, pueden utilizar prepruebas y pospruebas para analizar la evolución de los grupos antes y después del tratamiento experimental. Desde luego, no todos los diseños experimentales "puros" utilizan preprueba; aunque la posprueba sí es necesaria para determinar los efectos de las condiciones experimentales (Wiersma y Jurs, 2008). A continuación se muestran varios diseños experimentales "puros".

1. Diseño con posprueba únicamente y grupo de control

Este diseño incluye dos grupos: uno recibe el tratamiento experimental y el otro no (grupo de control). Es decir, la manipulación de la variable independiente alcanza sólo dos niveles: presencia y ausencia. Los sujetos se asignan a los grupos de manera aleatoria. Cuando concluye la manipulación, a ambos grupos se les administra una medición sobre la variable dependiente en estudio.

El diseño se diagrama de la siguiente manera:

$$\begin{array}{lll} RG_1 & X & 0_1 \\ RG_2 & — & 0_2 \end{array}$$

En este diseño, la única diferencia entre los grupos debe ser la presencia-ausencia de la variable independiente. Inicialmente son equivalentes y para asegurarse de que durante el experimento continúen siéndolo (salvo por la presencia o ausencia de dicha manipulación) el experimentador debe observar que no ocurra algo que sólo afecte a un grupo. La hora en que se efectúa el experimento debe ser la misma para ambos grupos (o ir mezclando un sujeto de un grupo con un sujeto del otro grupo, cuando la participación es individual), al igual que las condiciones ambientales y demás factores mencionados al hablar sobre la equivalencia de los grupos.

Wiersma y Jurs (2008) comentan que, de preferencia, la posprueba debe administrarse inmediatamente después de que concluya el experimento, en especial si la variable dependiente tiende a cambiar con el paso del tiempo. La posprueba se aplica de manera simultánea a ambos grupos.

La comparación entre las pospruebas de ambos grupos (0_1 y 0_2) nos indica si hubo o no efecto de la manipulación. Si ambas difieren significativamente[10] ($0_1 \neq 0_2$), esto nos indica que el tratamiento experimental tuvo un efecto a considerar. Por tanto, se acepta la hipótesis de diferencia de grupos. Si no hay diferencias ($0_1 = 0_2$), ello indica que no hubo un efecto significativo del tratamiento experimental (X). En este caso se acepta la hipótesis nula.

En ocasiones se espera que 0_1 sea mayor que 0_2. Por ejemplo, si el tratamiento experimental es un método educativo que facilita la autonomía por parte del alumno, y si el investigador formula la hipótesis de que incrementa el aprendizaje, cabe esperar que el nivel de aprendizaje del grupo experimental, expuesto a la autonomía, sea mayor que el nivel de aprendizaje del grupo de control, no expuesto a la autonomía: $0_1 > 0_2$.

En otras ocasiones se espera que 0_1 sea menor que 0_2. Por ejemplo, si el tratamiento experimental es un programa de televisión que supuestamente disminuye el prejuicio, el nivel de éste en el grupo experimental deberá ser menor que el del grupo de control: $0_1 < 0_2$. Pero si 0_1 y 0_2 son iguales, quiere decir que tal programa no reduce el prejuicio. Asimismo, puede suceder que los resultados vayan en contra de la hipótesis. Por ejemplo, en el caso del prejuicio, si 0_2 es menor que 0_1 (el nivel del prejuicio es menor en el grupo que no recibió el tratamiento experimental, esto es, el que no vio el programa televisivo).

Centro de recursos en línea

Las pruebas estadísticas que suelen utilizarse en este diseño y en otros que a continuación se revisarán, se incluyen en el capítulo 10, "Análisis de los datos cuantitativos", y en el capítulo 8 adicional del centro de recursos en línea del libro, "Análisis estadístico: segunda parte".

El diseño con posprueba únicamente y grupo de control puede extenderse para incluir más de dos grupos (tener varios niveles o modalidades de manipulación de la variable independiente). Los efectos de los tratamientos experimentales se investigan comparando las pospruebas de los grupos.

[10] Los estudiantes preguntan frecuentemente qué es una diferencia significativa. Si el promedio en la posprueba de un grupo en alguna variable es de 10 (por ejemplo), y en el otro es de 12, ¿esta diferencia es o no significativa? ¿Puede o no decirse que el tratamiento tuvo un efecto sobre la variable dependiente? A este respecto, hay pruebas o métodos estadísticos que indican si una diferencia entre dos o más cifras (promedios, porcentajes, puntuaciones totales, etc.) es o no significativa. Estas pruebas toman en cuenta aspectos como el tamaño de los grupos cuyos valores se comparan, las diferencias entre quienes integran los grupos y otros factores. Cada comparación entre grupos es distinta y ello lo consideran los métodos, los cuales se explicarán en el capítulo 10, "Análisis de los datos cuantitativos". No resultaría conveniente exponerlos aquí, porque habría que clarificar algunos aspectos estadísticos en los cuales se basan tales métodos, lo que provocaría confusión, sobre todo entre quienes se inician en el estudio de la investigación.

Su formato general sería:[11]

RG_1	X_1	0_1
RG_2	X_2	0_2
RG_3	X_3	0_3
•	•	•
•	•	•
•	•	•
RG_k	X_k	0_k
RG_{k+1}	—	0_{k+1}

En el diseño con posprueba únicamente y grupo de control, así como en sus posibles variaciones y extensiones, se logra controlar todas las fuentes de invalidación interna. La administración de pruebas no se presenta porque no hay preprueba. La inestabilidad no afecta porque los componentes del experimento son los mismos para todos los grupos (excepto la manipulación o los tratamientos experimentales), ni la instrumentación porque es la misma posprueba para todos, ni la maduración porque la asignación es al azar (si hay, por ejemplo, cinco sujetos en un grupo que se cansan fácilmente, habrá otros tantos en los otros grupos), ni la regresión estadística, porque si un grupo está regresando a su estado normal los otros también. La selección tampoco es problema, ya que si hay sujetos atípicos en un grupo, en los demás habrá igualmente sujetos atípicos. Todo se compensa. Las diferencias se pueden atribuir a la manipulación de la variable independiente y no a que los sujetos sean atípicos, pues la asignación aleatoria hace equivalentes a los grupos en este factor.

De este modo, si en los dos grupos sólo hubiera personas demasiado inteligentes y la variable independiente fuera el método de enseñanza, las diferencias en el aprendizaje se atribuirían al método y no a la inteligencia. La mortalidad no afecta, puesto que al ser los grupos equiparables, el número de personas que abandonen cada grupo tenderá a ser el mismo, salvo que las condiciones experimentales tengan algo en especial que haga que los sujetos abandonen el experimento; por ejemplo, que las condiciones sean amenazantes para los participantes, en cuyo caso la situación se detecta, analiza a fondo y corrige. De todas maneras el o la experimentadora tiene control sobre la situación, debido a que sabe que todo es igual para los grupos, con excepción del tratamiento experimental.

Otras interacciones tampoco pueden afectar los resultados, pues si la selección se controla, sus interacciones operarán de modo similar en todos los grupos. Además, la historia se controla si se vigila cuidadosamente que ningún acontecimiento afecte a un solo grupo. Y si ocurre el acontecimiento en todos los grupos, aunque afecte, lo hará de manera pareja.

En resumen, lo que influya en un grupo también influirá de manera equivalente en los demás. Este razonamiento se aplica a todos los diseños experimentales "puros".

Ejemplo

Diseño con posprueba únicamente, varios grupos y uno de control

Un investigador lleva a cabo un experimento para analizar cómo influye el estilo de liderazgo del supervisor en la productividad de los trabajadores.

Pregunta de investigación: ¿influye el estilo de liderazgo que ejerzan los supervisores de producción en una maquiladora sobre la productividad de los trabajadores de línea?

Hipótesis de investigación: "distintos estilos de liderazgo que ejerzan los supervisores tendrán diferentes efectos sobre la productividad".

[11] El factor "*k*" fue extraído de Wiersma y Jurs (2008) e indica "un número tal de grupos". Otros autores utilizan "*n*". En los ejemplos, tal factor implica el número del último grupo con tratamiento experimental más uno. Desde luego, el grupo de control se incluye al final y el número que le corresponde a su posprueba será el último.

En una planta maquiladora, se asigna al azar a 90 trabajadores de línea a tres condiciones experimentales: 1) 30 realizan una tarea bajo el mando de un supervisor con estilo autocrático, 2) 30 ejecutan la tarea bajo el mando de un supervisor con estilo democrático y 3) 30 efectúan la tarea bajo el mando de un supervisor con estilo liberal (que no supervisa directamente, no ejerce presión y es permisivo). Por último, 30 más son asignados en forma aleatoria al grupo de control donde no hay supervisor. En total, son 120 trabajadores.

Se forman grupos de 10 trabajadores para el desempeño de la tarea (armar un sistema de arneses o cables para vehículos automotores). Por tanto, habrá 12 grupos de trabajo repartidos en tres tratamientos experimentales y un grupo de control. La tarea es la misma para todos y los instrumentos de trabajo también, al igual que el ambiente físico (iluminación, temperatura, etc.). Las instrucciones son uniformes.

Se ha preparado a tres supervisores (desconocidos para todos los participantes) para que ejerzan los tres estilos (democrático, autocrático y liberal). Los supervisores se distribuyen al azar entre los horarios.

Supervisor	Estilos		
	Autocrático	Democrático	Liberal
Supervisor 1 trabaja con...	10 sujetos 10:00-14:00 h Lunes	10 sujetos 15:00-19:00 h Lunes	10 sujetos 10:00-14:00 h Martes
Supervisor 2 trabaja con...	10 sujetos 15:00-19:00 h Lunes	10 sujetos 10:00-14:00 h Martes	10 sujetos 10:00-14:00 h Lunes
Supervisor 3 trabaja con...	10 sujetos 10:00-14:00 h Martes	10 sujetos 10:00-14:00 h Lunes	10 sujetos 15:00-19:00 h Lunes
Sin supervisor	10 sujetos 10:00-14:00 h Lunes	10 sujetos 15:00-19:00 h Lunes	10 sujetos 10:00-14:00 h Martes

Si se observa, los tres supervisores interactúan en todas las condiciones (siguen los tres estilos), con el propósito de evitar que la apariencia física o la personalidad del supervisor afecte solamente a una de ellas. Es decir, si un supervisor es más "carismático" que los demás e impacta la productividad, influirá en los tres grupos por igual.

El horario está controlado, puesto que los tres estilos se aplican en todas las horas en que se lleva a cabo el experimento. Es decir, las tres condiciones siempre se realizan en forma simultánea. Este ejemplo se esquematizaría de la siguiente manera:

RG_1 X_1 (supervisión con estilo autocrático) 0_1
RG_2 X_2 (supervisión con estilo democrático) 0_2
RG_3 X_3 (supervisión con estilo liberal) 0_3
RG_4 — (sin supervisión) 0_4

} Comparaciones en productividad

Ejemplo

En el ejemplo del selenio que se comentó previamente en esta obra, además de ponderar la literatura podría realizarse un experimento o prueba clínica aplicando este diseño a, digamos, tres grupos de mujeres con tumores cancerígenos: a uno que se le administrara durante un año cierto complemento alimenticio de selenio en cápsulas (por ejemplo, 200 µg diarios), a otro solamente 100 µg y a un tercero no se le suministrara *Se* (de control). La posprueba consistiría en evaluar si el tratamiento reduce el crecimiento de los tumores cancerígenos en aquellas pacientes que se encuentran en la etapa inicial de la enfermedad.

Desde luego, se controlarían posibles fuentes de invalidación o contaminación como la dieta (que la alimentación sea la misma para todas los participantes, porque hay alimentos que contienen selenio, como el pescado y el huevo), y la asignación al azar igualaría a los grupos en edad, región geográfica donde viven (vinculada a la dieta), nivel socioeconómico y otras variables que pudieran afectar.

Una cuestión que debe valorarse en esta clase de intervenciones es que diversos estudios han demostrado que administrar selenio puede ser muy delicado, ya que altos niveles de este cofactor esencial en los sistemas antioxidantes endógenos más importantes del cuerpo humano pueden tener efectos en la salud como el riesgo potencial de desarrollar diabetes tipo 2 (Muecke *et al.,* 2010), además de otros efectos secundarios. Asimismo, un protocolo de tal naturaleza tiene que ser sometido a distintas instancias de ética médica y comunidades científicas. El investigador debe asegurarse de que el selenio no vaya a tener efectos perjudiciales en las participantes (no sólo con la vigilancia del crecimiento de los tumores, sino realizando también otras pruebas clínicas para evaluar permanentemente el estado de salud y suspender el experimento a la mínima sospecha de otros efectos). Y también existe el serio dilema de las enfermas del grupo de control, pues se les niega la posibilidad de mejorar al no suministrarles el selenio, por esto, tal vez la opción sería no tener grupo testigo o que éste se encuentre constituido por mujeres que ya hayan fallecido y se posea información sobre la evolución de sus tumores durante la enfermedad, además de que posean un perfil similar a las participantes del experimento (emparejarlas). En este caso, se mezclarían asignación al azar y emparejamiento en la constitución de los grupos. Roberto Hernández-Sampieri recomienda a los estudiantes discutir las cuestiones éticas de la experimentación con su profesor o profesora de métodos de investigación.[12]

2. Diseño con preprueba-posprueba y grupo de control

Este diseño incorpora la administración de prepruebas a los grupos que componen el experimento. Los participantes se asignan al azar a los grupos y después se les aplica simultáneamente la preprueba; un grupo recibe el tratamiento experimental y otro no (es el grupo de control); por último, se les administra, también simultáneamente, una posprueba (Petrosko, 2004). El diseño se diagrama como sigue:

$$\begin{array}{cccc} RG_1 & 0_1 & X & 0_2 \\ RG_2 & 0_3 & — & 0_4 \end{array}$$

La adición de la prueba previa ofrece dos ventajas: primera, sus puntuaciones sirven para fines de control en el experimento, pues al compararse las prepruebas de los grupos se evalúa qué tan adecuada fue la asignación aleatoria, lo cual es conveniente con grupos pequeños. En grupos grandes, la técnica de distribución aleatoria funciona, pero cuando tenemos grupos de 15 personas no está de más evaluar qué tanto funcionó la asignación al azar. La segunda ventaja reside en que es posible analizar el puntaje-ganancia de cada grupo (la diferencia entre las puntuaciones de la preprueba y la posprueba).

El diseño elimina el impacto de todas las fuentes de invalidación interna por las mismas razones que se argumentaron en el diseño anterior (diseño con posprueba únicamente y grupo de control). Y la administración de pruebas queda controlada, ya que si la preprueba afecta las puntuaciones de la posprueba lo hará de manera similar en ambos grupos. Lo que influye en un grupo deberá afectar de la misma manera en el otro, para mantener la equivalencia entre ambos.

En algunos casos, para no repetir exactamente la misma prueba, se desarrollan dos versiones equivalentes (que produzcan los mismos resultados).[13] La historia se controla viendo que ningún acontecimiento afecte a sólo un grupo.

[12] Aunque este diseño más bien podría ser una serie cronológica experimental (que se comentará más adelante), porque habría múltiples aplicaciones del estímulo y pospruebas (mamografías comparativas o modelamiento de una función continua del tamaño del tumor que considera volumen, diámetro y tiempo, así como edad y datos de la población donde se efectúa el estudio —tasa de crecimiento—) (Weedon-Fekjær *et al.*, 2008 y Muecke *et al.*, 2010).

[13] Hay procedimientos para obtener pruebas "paralelas" o "gemelas", los cuales se comentan en el capítulo 9. Si no se asegura la equivalencia de las pruebas, no se pueden comparar las puntuaciones producidas por ambas. Es decir, se pueden presentar las fuentes de invalidación interna: "inestabilidad", "instrumentación" y "regresión estadística".

www.elosopanda.com | jamespoetrodriguez.com

Es posible extender este diseño para incluir más de dos grupos, lo cual se diagramaría de una manera general del siguiente modo:

$$
\begin{array}{cccc}
RG_1 & 0_1 & X_1 & 0_2 \\
RG_2 & 0_3 & X_2 & 0_4 \\
RG_3 & 0_5 & X_3 & 0_6 \\
\bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet \\
RG_k & 0_{2k-1} & Xk & 0_{2k} \\
RG_{k+1} & 0_{2k+1} & — & 0_{2(k+1)}
\end{array}
$$

Se tienen diversos tratamientos experimentales y un grupo de control. Si éste es excluido, el diseño se llamaría "diseño de preprueba-posprueba con grupos distribuidos aleatoriamente" (Simon, 1985).

Ejemplo

Diseño de preprueba-posprueba con grupo de control

Un investigador desea analizar el efecto de utilizar un DVD didáctico con canciones para enseñar hábitos higiénicos a los niños de cuatro a cinco años de edad.

Pregunta de investigación: ¿los DVD didácticos musicalizados son más eficaces para enseñar hábitos higiénicos a los niños de cuatro a cinco años de edad que otros métodos tradicionales de enseñanza?

Hipótesis de investigación: "los DVD didácticos constituyen un método más eficaz de enseñanza de hábitos higiénicos a niños de cuatro a cinco años, que la explicación verbal y los libros impresos".

Cien niños de cuatro a cinco años de edad se asignan al azar a cuatro grupos:

1. Un grupo recibirá instrucción sobre hábitos higiénicos por medio de un DVD con caricaturas y canciones, con duración de 30 minutos.
2. Otro grupo recibirá explicaciones de hábitos higiénicos de una maestra instruida para ello, la ilustración durará 30 minutos y no se permiten preguntas.
3. El tercer grupo leerá un libro infantil ilustrado con explicaciones sobre hábitos higiénicos (la publicación está diseñada para que un niño promedio de cuatro a cinco años la lea en 30 minutos).
4. El grupo de control verá un DVD sobre otro tema durante 30 minutos.

Los grupos permanecerán simultáneamente en cuatro salones de clases. Todas las explicaciones (DVD, instrucción oral y libro) contendrán la misma información y las instrucciones son estándares.

Antes del inicio del tratamiento experimental, a todos los grupos se les aplicará una prueba sobre conocimiento de hábitos higiénicos especialmente diseñada para niños, del mismo modo se aplicará una vez que hayan recibido la explicación por el medio que les correspondió. El ejemplo se esquematizaría de la forma en que lo muestra la tabla 7.2.

Tabla 7.2 Diagrama del ejemplo de diseño de preprueba-posprueba con grupo de control

RG_1	0_1	Video didáctico (X_1)	0_2
RG_2	0_3	Explicación verbal (X_2)	0_4
RG_3	0_5	Lectura de libro ilustrado (X_3)	0_6
RG_4	0_7	No estímulo	0_8
	↑ Prueba de conocimientos higiénicos		↑ Prueba de conocimientos higiénicos

Las posibles comparaciones en este diseño son: *a*) las prepruebas entre sí (0_1, 0_3, 0_5 y 0_7), *b*) las pospruebas entre sí para analizar cuál fue el método de enseñanza más efectivo (0_2, 0_4, 0_6 y 0_8), *c*) el puntaje-ganancia de cada grupo (0_1 frente a 0_2, 0_3 vs. 0_4, 0_5 frente a 0_6 y 0_7 vs. a 0_8), y *d*) los puntajes ganancia de los grupos entre sí. Al igual que en todos los diseños experimentales, es posible tener más de una variable dependiente (por ejemplo, interés por los hábitos higiénicos, disfrute del método de enseñanza, etc.). En este caso, las prepruebas y pospruebas medirán diversas variables dependientes.

Veamos algunos posibles resultados de este ejemplo y sus interpretaciones:

1. Resultado: $0_1 \neq 0_2$, $0_3 \neq 0_4$, $0_5 \neq 0_6$, $0_7 \neq 0_8$; pero $0_2 \neq 0_4$, $0_2 \neq 0_6$, $0_4 \neq 0_6$.
   Interpretación: hay efectos de todos los tratamientos experimentales, incluso en el grupo testigo, pero son diferentes.
2. Resultado: $0_1 = 0_3 = 0_5 = 0_2 = 0_6 = 0_7 = 0_8$; pero $0_3 \neq 0_4$.
   Interpretación: no hay efectos de X_1 ni X_3, pero sí hay efectos de X_2.
3. Resultado: $0_1 = 0_3 = 0_5 = 0_7$ y $0_2 = 0_4 = 0_6 = 0_8$; pero 0_1, 0_3, 0_5 y $0_7 < 0_2$, 0_4, 0_6 y 0_8.
   Interpretación: no hay efectos de los tratamientos experimentales, sino un posible efecto de sensibilización de la preprueba o de maduración en todos los grupos (éste es parejo y se encuentra bajo control).

3. Diseño de cuatro grupos de Solomon

Solomon (1949) propuso un diseño que era la mezcla de los dos anteriores, con dos grupos experimentales y dos de control. Sólo a uno de los grupos experimentales y a uno de los grupos de control se les administra la preprueba; a los cuatro grupos se les aplica la posprueba. Los participantes se asignan en forma aleatoria.

El diseño se diagrama así:

RG_1	0_1	X	0_2
RG_2	0_3	—	0_4
RG_3	—	X	0_5
RG_4	—	—	0_6

La ventaja de este diseño es que el experimentador tiene la posibilidad de verificar los posibles efectos de la preprueba sobre la posprueba, puesto que a unos grupos se les administra un test previo y a otros no. El diseño de Solomon se amplía en el capítulo 5 adicional del centro de recursos en línea, Diseños: segunda parte.

Centro de recursos en línea

4. Diseños experimentales de series cronológicas múltiples

Los tres diseños experimentales que se han comentado sirven más bien para analizar efectos inmediatos o a corto plazo. En ocasiones, el experimentador está interesado en analizar efectos en el mediano o largo plazo, porque tiene bases para suponer que la influencia de la variable independiente sobre la dependiente tarda en manifestarse. Por ejemplo, programas de difusión de innovaciones, métodos educativos, modelos de entrenamiento, procesos de mejora de la calidad o estrategias de las psicoterapias.

Centro de recursos en línea

Asimismo, en otras situaciones se busca evaluar la evolución del efecto en el corto, mediano y largo plazos (no solamente el resultado). También, en ocasiones la aplicación única del estímulo no tiene efectos (una dosis de un medicamento o sustancia —como en el caso del selenio—, un único programa televisivo, unos cuantos anuncios en la radio, etc.). En tales casos es conveniente adoptar diseños con varias pospruebas o bien con diversas prepruebas y pospruebas, con repetición del estímulo, con varios tratamientos aplicados a un mismo grupo y otras condiciones. A estos diseños se les conoce como **series cronológicas experimentales** y son explicados en el capítulo 5 adicional del centro de recursos en línea: "Diseños: segunda parte". En realidad el término "serie cronológica" se aplica a cualquier diseño en el que se efectúe al paso del tiempo varias observaciones o mediciones sobre una o más variables, sea o no experimental, sólo que en este caso se les llama experimentales porque reúnen los requisitos para serlo. En estos diseños se pueden tener dos o más grupos y los participantes se asignan al azar.

Serie cronológica Diseño en el que se efectúan al paso del tiempo varias observaciones o mediciones sobre una o más variables, sea o no experimental.

5. Diseños factoriales

En ocasiones, el investigador pretende analizar experimentalmente el efecto que sobre las variables dependientes tiene la manipulación de más de una variable independiente. Por ejemplo, analizar el efecto que poseen sobre la productividad de los trabajadores: 1) la fuente de realimentación sobre el desempeño en el trabajo (de parte del supervisor en persona, por escrito y por medio de los compañeros) y 2) el tipo de realimentación (positiva, negativa, y positiva y negativa). En este caso se manipulan dos variables independientes.

Veamos en otro ejemplo: determinar el efecto de tres medicamentos distintos (primera variable independiente, clase de medicamento) y la dosis diaria (segunda variable independiente, con dos niveles, supongamos 40 mg y 20 mg) sobre la cura de una enfermedad (variable dependiente). También aquí tenemos dos independientes. Pero podríamos tener tres o más: conocer cómo afectan en el nivel de aceleración de un vehículo (dependiente), el peso del chasís (dos diferentes pesos), el material con que está fabricado (supongamos tres tipos de materiales), el tamaño del rin de las ruedas (14, 15 y 16 pulgadas) y el diseño de la carrocería (por ejemplo, dos diseños distintos). Cuatro variables independientes. Estos diseños se conocen como factoriales.

Los **diseños factoriales** manipulan dos o más variables independientes e incluyen dos o más niveles o modalidades de presencia en cada una de las variables independientes. Se utilizan muy a menudo en la investigación experimental. La preparación básica de un diseño factorial consiste en que todos los niveles o modalidades de cada variable independiente son tomados en combinación con todos los niveles o modalidades de las otras variables independientes (Babbie, 2014 y Wiersma y Jurs, 2008). Tales diseños se exponen y evalúan en el capítulo 5 adicional del centro de recursos en línea, "Diseños: segunda parte".

Centro de recursos en línea

Ejemplo

Un ejemplo real es el estudio de Pérez, Arango y Agudelo (2009) para determinar el efecto que tienen los factores experiencia del operario, tipo de dobladora, clase de material utilizado y grosor sobre la longitud y el ángulo de doblado de las piezas de metal producidas (estas dos últimas, las variables dependientes). Su diseño fue experimental factorial, ya que se consideraron estas cuatro variables independientes con dos grados o modalidades cada una: experiencia del operario (menos de cinco años/más de cinco años), tipo de dobladora (hidráulica electrónica/que conduce por aire), clase de material utilizado (acero laminado en frío/acero galvanizado) y su grosor (0.8 mm/1 mm). Se realizaron dos réplicas para cada combinación m = 1, 2 n = 2. Los autores consideraron conveniente el uso de un diseño 4 × 2, debido a que proporciona el menor número de corridas experimentales en las que las cuatro variables "pueden estudiarse en un diseño factorial completo. Además este tipo de diseño reduce los costos de la experimentación" (Pérez, Arango y Agudelo, 2009, p. 149).

¿Qué es la validez externa?

Un experimento debe buscar, ante todo, *validez interna*, es decir, confianza en los resultados. Si no se logra, no hay experimento "puro". Lo primero es eliminar las fuentes que atentan contra dicha validez. Pero la validez interna es sólo una parte de la validez de un experimento; además, es muy deseable que el experimento tenga validez externa. La **validez externa** se refiere a qué tan generalizables son los resultados de un experimento a situaciones no experimentales, así como a otros participantes o poblaciones. Responde a la pregunta: ¿lo que encontré en el experimento a qué tipos de personas, grupos, fenómenos, contextos y situaciones se aplica?

Validez externa Posibilidad de generalizar los resultados de un experimento a situaciones no experimentales, así como a otras personas, casos y poblaciones.

Por ejemplo, si hacemos un experimento con métodos de aprendizaje y los resultados se pueden generalizar a la enseñanza cotidiana en las escuelas de educación elemental (primaria) del país, el experimento tendrá validez externa; del mismo modo, si se generalizan a la enseñanza cotidiana de nivel infantil, elemental y secundaria (media), tendrá aún mayor validez externa.

Así, los resultados de experimentos sobre liderazgo y motivación que se extrapolen a situaciones diarias de trabajo en las empresas, la actividad de las organizaciones gubernamentales y no guberna-

mentales, incluso al funcionamiento de los grupos de niños y jóvenes exploradores (*boy scouts*), son experimentos con validez externa.

Fuentes de invalidación externa

Hay diversos factores que llegan a amenazar la validez externa; los más comunes se mencionan en la tabla 7.3 (una explicación más detallada, así como ejemplos y otras fuentes potenciales, las podrá encontrar el lector en el capítulo 5 adicional, "Diseños: segunda parte", que se puede descargar en el centro de recursos en línea).

Tabla 7.3 Principales fuentes de invalidación externa[14]

Fuente o amenaza a la validez externa	Descripción de la amenaza
Efecto reactivo o de interacción de las pruebas	Cuando la preprueba aumenta o disminuye la sensibilidad o reacción de los participantes a la variable experimental y los resultados obtenidos para una población con preprueba no pueden generalizarse a quienes forman parte de esa población pero sin preprueba.
Efecto de interacción entre los errores de selección y el tratamiento experimental	Elegir personas con una o varias características que hagan que el tratamiento experimental produzca un efecto, que no se daría si las personas no tuvieran esas características. A veces este factor se presenta cuando se reclutan voluntarios para la realización de los experimentos.
Efectos reactivos de los tratamientos (Hawthorne)	"Artificialidad" de las condiciones que puede hacer que el contexto experimental resulte atípico respecto a la manera en que se aplica regularmente el tratamiento.
Interferencia de tratamientos múltiples	Que algunos tratamientos nulifiquen el efecto de otros.
Imposibilidad de replicar los tratamientos	Cuando los tratamientos son tan complejos que no pueden replicarse en situaciones no experimentales.
Efectos de novedad e interrupción	Un nuevo tratamiento puede tener resultados positivos simplemente por ser percibido como novedoso, o bien, lo contrario: tener un efecto negativo porque interrumpe las actividades normales de los participantes.
El experimentador	Que genere alteraciones o cambios que no se presentan en situaciones no experimentales. Es decir que el tratamiento solamente tenga efecto con la intervención del experimentador.
Interacción entre la historia o el lugar y los efectos del tratamiento experimental	Imposibilidad de duplicar un experimento conducido en un contexto en particular (tiempo y lugar) o que los resultados del experimento no pueden generalizarse a otros lugares o ambientes.

Para lograr una mayor validez externa es conveniente tener casos o grupos lo más parecidos posible a la mayoría de las personas o poblaciones a las cuales se desea generalizar, y repetir el experimento varias veces con diferentes grupos o en distintos ambientes (hasta donde el presupuesto y el tiempo lo permitan). También, desde luego, tratar de que el contexto experimental sea lo más similar al contexto específico que se pretende generalizar. Por ejemplo, si se trata de métodos de enseñanza resultaría muy conveniente que se usen aulas similares a las que normalmente utilizan los participantes y que las instrucciones las proporcionen los maestros de siempre. Claro que a veces no es factible; sin embargo, el experimentador debe esforzarse para que quienes participan no sientan, o que sea lo menos posible, que están experimentando con ellos.

[14] Basada en Hernández-Sampieri *et al.* (2013), Mertens (2010) y Campbell (1975).

¿Cuáles pueden ser los contextos generales de los experimentos?

En la literatura sobre la investigación se distinguen dos contextos generales en los que llega a tomar lugar un diseño experimental: laboratorio y campo. Así, se habla de **experimentos de laboratorio** y **experimentos de campo**.

Experimentos de laboratorio Experimento en que el efecto de todas o casi todas las variables independientes influyentes no concernientes al problema de investigación se mantiene reducido lo más posible.

Experimentos de campo Experimento en una situación más real o natural en la que el investigador manipula una o más variables.

Los **experimentos de laboratorio** se realizan en condiciones controladas, en las cuales el efecto de las fuentes de invalidación interna es eliminado, así como el de otras posibles variables independientes que no son manipuladas o no interesan (Hernández-Sampieri *et al.,* 2013 y Crano, 2003). Los **experimentos de campo** son estudios efectuados en una situación "realista" en la que el investigador manipula una o más variables independientes en condiciones tan cuidadosamente controladas como lo permite la situación (Gerber y Green, 2012; Smith, 2004 y Kerlinger y Lee, 2002).

La diferencia esencial entre ambos contextos generales es el "realismo" con que los experimentos se llevan a cabo, es decir, el grado en que el ambiente es natural para los sujetos.

Por ejemplo, si creamos salas para ver televisión y las acondicionamos de tal modo que se controle el ruido exterior, la temperatura y otros distractores; incluimos equipo de filmación oculto y llevamos a los niños para que vean programas de televisión grabados. De esta manera estamos realizando un experimento de laboratorio (situación construida "artificialmente"). En cambio, si el experimento se lleva a cabo en el ambiente cotidiano de los sujetos (como en sus casas), se trata de un experimento de campo.

Los experimentos de laboratorio generalmente logran un control más riguroso que los experimentos de campo (Festinger, 1993), pero estos últimos suelen tener mayor validez externa. Ambos tipos de experimento son deseables.

En ciencias sociales, algunos autores han acusado a los experimentos de laboratorio de "artificialidad", de mantener distancia respecto al grupo estudiado, de imposibilitar un entendimiento completo del fenómeno que se analiza, de ser reduccionistas y de que descontextualizan la conducta humana para simplificar su interpretación (Mertens, 2010). Sin embargo, como argumenta Festinger (1993, p. 139):

> Esta crítica requiere ser evaluada, pues probablemente sea consecuencia de una equivocada interpretación de los fines del experimento de laboratorio. Un experimento de laboratorio no necesita, y no debe, constituir un intento de duplicar una situación de la vida real. Si se quisiera estudiar algo en una situación de este tipo, sería bastante tonto tomarse el trabajo de organizar un experimento de laboratorio para reproducir dicha situación. ¿Por qué no estudiarla directamente? El experimento de laboratorio debe tratar de crear una situación en la cual se vea claramente cómo operan las variables en situaciones especialmente identificadas y definidas. El hecho de que pueda encontrarse o no tal situación en la vida real no tiene importancia. Evidentemente, nunca puede encontrarse en la vida real la situación de la mayor parte de los experimentos de laboratorio. No obstante, en el laboratorio podemos determinar con exactitud en qué medida una variable específica afecta la conducta o actitudes en condiciones especiales o puras.

¿Qué alcance tienen los experimentos y cuál es el enfoque del que se derivan?

Debido a que analizan las relaciones entre una o más variables independientes y una o más dependientes, así como los efectos causales de las primeras sobre las segundas, son estudios explicativos (que obviamente determinan correlaciones). Se trata de diseños que se fundamentan en el enfoque cuantitativo y en el paradigma deductivo. Se basan en hipótesis preestablecidas, miden variables y su aplicación debe sujetarse al diseño concebido con antelación; al desarrollarse, el investigador está centrado en la validez, el rigor y el control de la situación de investigación. Asimismo, el análisis estadístico

www.elosopanda.com | jamespoetrodriguez.com

resulta fundamental para lograr los objetivos de conocimiento. Como señalan Feuer, Towne y Shavelson (2002), su fin es estimar efectos causales.

Simbología de los diseños con emparejamiento en lugar de asignación al azar

Como ya se comentó, otra técnica para hacer inicialmente equivalentes a los grupos es el emparejamiento. Desde luego, este método es menos preciso que la asignación al azar. Los diseños se representan con una "*E*" de emparejamiento, en lugar de la "*R*" (asignación aleatoria o al azar). Por ejemplo,

$$\begin{array}{llll} E & G_1 & X_1 & 0_1 \\ E & G_2 & X_2 & 0_2 \\ E & G_3 & — & 0_3 \end{array}$$

¿Qué otros experimentos hay?: cuasiexperimentos

Los **diseños cuasiexperimentales** también manipulan deliberadamente, al menos, una variable independiente para observar su efecto sobre una o más variables dependientes, sólo que difieren de los experimentos "puros" en el grado de seguridad que pueda tenerse sobre la equivalencia inicial de los grupos. En los diseños cuasiexperimentales, los sujetos no se asignan al azar a los grupos ni se emparejan, sino que dichos grupos ya están conformados antes del experimento: son grupos intactos (la razón por la que surgen y la manera como se integraron es independiente o aparte del experimento). Por ejemplo, si fueran tres grupos escolares formados con anterioridad a la realización del experimento, y cada uno de ellos constituye un grupo experimental. Veámoslo gráficamente:

Grupo *A* (30 estudiantes)	Grupo experimental con X_1
Grupo *B* (26 estudiantes)	Grupo experimental con X_2
Grupo *C* (34 estudiantes)	Grupo de control

Otros ejemplos serían utilizar grupos terapéuticos ya integrados, equipos deportivos constituidos, trabajadores de turnos establecidos o grupos de habitantes de distintas regiones geográficas (que ya estén agrupados por zona).

Los diseños cuasiexperimentales específicos se revisan en el capítulo 5 adicional del centro de recursos en línea: "Diseños: segunda parte".

Centro de recursos en línea

Pasos de un experimento

A continuación mencionamos los principales pasos que se dan en el desarrollo de un experimento:

Paso 1: Decidir cuántas variables independientes y dependientes deberán incluirse. No necesariamente el mejor experimento es el que incluye el mayor número de variables; deben incluirse las variables que sean necesarias para probar las hipótesis, alcanzar los objetivos y responder las preguntas de investigación.

Paso 2: Elegir los niveles o modalidades de manipulación de las variables independientes y traducirlos en tratamientos experimentales.

Paso 3: Desarrollar el instrumento o instrumentos para medir las variables dependientes.

Paso 4: Seleccionar una muestra de casos o personas del tipo o perfil que nos interesa.

Paso 5: En el caso de que sean individuos, reclutarlos. Esto implica ponerse en contacto con ellos, darles las explicaciones necesarias, obtener su consentimiento e indicarles lugar, día, hora y persona con quien deben presentarse. Es conveniente proporcionarles facilidades para que acudan al experimento (si se les puede brindar transporte, entregarles un mapa con los señalamientos precisos, etc.). También hay que darles cartas (a ellos o alguna institución a la que pertenezcan para facilitar el apoyo; por ejemplo, en escuelas

a los directivos, maestros y padres de familia) y recordarles su participación el día anterior a la realización del experimento.

Las personas deben sentirse motivadas para participar. Por tanto, resulta muy conveniente darles algún regalo atractivo (a veces, simbólico). Por ejemplo, a amas de casa, una canasta de productos básicos; a ejecutivos, una canasta con dos o tres artículos; a estudiantes, créditos escolares, etc., además de expedirles una carta de agradecimiento.

Paso 6: Seleccionar el diseño experimental o cuasiexperimental apropiado para nuestras hipótesis, objetivos y preguntas de investigación.

Paso 7: Planear cómo vamos a manejar los casos o a los participantes. Con personas, elaborar una ruta crítica sobre qué van a hacer desde que llegan al lugar del experimento hasta que se retiran.

Paso 8: En el caso de experimentos "puros", dividirlos al azar o emparejarlos; y en el caso de cuasiexperimentos, analizar cuidadosamente las propiedades de los grupos intactos.

Paso 9: Aplicar las prepruebas (cuando las haya), los tratamientos y las pospruebas.

Asimismo, resulta conveniente tomar nota del desarrollo del experimento, llevar una bitácora minuciosa de todo lo ocurrido.

En los últimos años algunos autores señalan que, por razones éticas, el estímulo o tratamiento experimental debe ser discutido con los sujetos antes de aplicarlo (Mertens, 2010), sobre todo si exige esfuerzo físico o puede tener un fuerte impacto emocional. Esto es adecuado, siempre y cuando no se convierta en una fuente de invalidación interna o de anulación del experimento. Asimismo, se recomienda que si por medio del tratamiento se beneficia a un grupo (por ejemplo, con un método educativo o un curso), una vez concluido el experimento se administre a los demás grupos, para que también gocen de sus beneficios.

Centro de recursos en línea

En el capítulo 5 adicional del centro de recursos en línea, "Diseños: segunda parte" también se explica cómo controlar la influencia de variables intervinientes y otros temas importantes.

Diseños no experimentales

¿Qué es la investigación no experimental cuantitativa?

Podría definirse como la investigación que se realiza sin manipular deliberadamente variables. Es decir, se trata de estudios en los que **no** hacemos variar en forma intencional las variables independientes para ver su efecto sobre otras variables. Lo que hacemos en la **investigación no experimental** es observar fenómenos tal como se dan en su contexto natural, para analizarlos (The SAGE Glossary of the Social and Behavioral Sciences, 2009b).

Investigación no experimental Estudios que se realizan sin la manipulación deliberada de variables y en los que sólo se observan los fenómenos en su ambiente natural para analizarlos.

En un experimento, el investigador prepara deliberadamente una situación a la que son expuestos varios casos o individuos. Esta situación consiste en recibir un tratamiento, una condición o un estímulo en determinadas circunstancias, para después evaluar los efectos de la exposición o aplicación de dicho tratamiento o tal condición. Por decirlo de alguna manera, en un experimento se "construye" una realidad.

En cambio, en un estudio no experimental no se genera ninguna situación, sino que se observan situaciones ya existentes, no provocadas intencionalmente en la investigación por quien la realiza. En la investigación no experimental las variables independientes ocurren y no es posible manipularlas, no se tiene control directo sobre dichas variables ni se puede influir en ellas, porque ya sucedieron, al igual que sus efectos.

La investigación no experimental es un parteaguas de varios estudios cuantitativos, como las encuestas de opinión, los estudios *ex post-facto* retrospectivos y prospectivos, etc. Para ilustrar la diferencia entre un estudio experimental y uno no experimental consideremos el siguiente ejemplo. Claro está que no sería ético un experimento que obligara a las personas a consumir una bebida que afecta gravemente la salud. El ejemplo es sólo para ilustrar lo expuesto y quizá parezca un tanto burdo, pero es ilustrativo.

Ejemplo

Para esclarecer la diferencia entre la investigación experimental y la investigación no experimental

Vamos a suponer que un investigador desea analizar el efecto que produce el consumo de alcohol sobre los reflejos humanos. Su hipótesis es: "a mayor consumo de alcohol, mayor lentitud en los reflejos de las personas". Si decidiera seguir un enfoque experimental, asignaría al azar los sujetos a varios grupos. Supóngase cuatro grupos: un primer grupo donde los participantes ingirieran un elevado consumo de alcohol (siete copas de tequila o brandy), un segundo grupo que tuviera un consumo medio de alcohol (cuatro copas), un tercer grupo que bebiera un consumo bajo de alcohol (una sola copa) y un cuarto grupo de control que no ingiriera nada de alcohol. Controlaría el lapso en el que todos los sujetos consumen su "ración" de alcohol (copa o copas), así como otros factores (misma bebida, cantidad de alcohol servida en cada copa, etc.). Finalmente, mediría la calidad de la respuesta de los reflejos en cada participante y compararía los grupos, para determinar el efecto del consumo de alcohol sobre los reflejos humanos, así probaría o refutaría su hipótesis.

Desde luego, el enfoque podría ser cuasiexperimental (grupos intactos) o asignar los sujetos a los grupos por emparejamiento (digamos en cuanto al género, que influye en la resistencia al alcohol, pues la mayoría de las mujeres suelen tolerar menos cantidades que los hombres).

Por el contrario, si decidiera seguir un enfoque no experimental, el investigador podría acudir a lugares donde se localicen distintas personas con diferentes consumos de alcohol (por ejemplo, oficinas donde se haga la prueba del nivel de consumo de alcohol, como una estación de policía). Encontraría a personas que han bebido cantidades elevadas, medias y bajas de alcohol, así como a quienes no lo han ingerido. Mediría la calidad de sus reflejos, llevaría a cabo sus comparaciones y establecería el efecto del consumo de alcohol sobre los reflejos humanos, analizando si aporta evidencia en favor o en contra de su hipótesis.

En un estudio experimental se construye el contexto y se manipula de manera intencional la variable independiente (en este caso, el consumo del alcohol) y se observa el efecto de esta manipulación sobre la variable dependiente (aquí, la calidad de los reflejos). Es decir, el investigador influyó directamente en el grado de consumo de alcohol de los participantes. En la investigación no experimental no hay ni manipulación intencional ni asignación al azar. Los sujetos ya habían consumido cierto nivel de alcohol y en este hecho el investigador no tuvo nada que ver: no influyó en la cantidad de consumo de alcohol de los participantes. Era una situación que ya existía, ajena al control directo que hay en un experimento. En la investigación no experimental se eligieron personas con diferentes niveles de consumo, los cuales se generaron por muchas causas, pero no por la manipulación intencional y previa del consumo de alcohol. En resumen, en un estudio no experimental los individuos ya pertenecían a un grupo o nivel determinado de la variable independiente por autoselección.

Esta diferencia esencial genera características distintas en la investigación experimental y la no experimental, que serán comentadas cuando se comparen ambos enfoques. Para ello es necesario profundizar en los tipos de investigación no experimental.

La investigación experimental tiene alcances iniciales y finales correlacionales y explicativos. La investigación **no** experimental es sistemática y empírica en la que las variables independientes no se manipulan porque ya han sucedido. Las inferencias sobre las relaciones entre variables se realizan sin intervención o influencia directa, y dichas relaciones se observan tal como se han dado en su contexto natural.

Un ejemplo no científico (y tal vez demasiado coloquial) para abundar en la diferencia entre un experimento y un **no** experimento serían las siguientes situaciones:

Experimento	Hacer enojar intencionalmente a una persona y ver sus reacciones.
No experimento	Ver las reacciones de esa persona cuando llega enojada.

Mertens (2010) señala que la investigación no experimental es apropiada para variables que no pueden o deben ser manipuladas o resulta complicado hacerlo. Algunos ejemplos se muestran en la tabla 7.4.

Tabla 7.4 Variables no manipulables o difícilmente manipulables en experimentos, y apropiadas más bien para estudios no experimentales

Tipos	Ejemplos
Características inherentes de personas u objetos que son complejas de manipular.	Hábitat de un animal, fuertes incrementos salariales, antigüedad en el trabajo...
Características que no pueden ser manipuladas por razones éticas.	Consumo de alcohol, tabaco o un medicamento (si la persona se encuentra saludable), agresiones físicas, adopción, estado civil de los padres (divorciados,casados, unión libre, etc.), impedimentos físicos...
Características que no es posible manipular.	Personalidad (todos sus rasgos), energía explosiva de un volcán, hechos históricos pasados, masa de un meteorito...

¿Cuáles son los tipos de diseños no experimentales?

Distintos autores han adoptado diversos criterios para catalogar la investigación no experimental. En este libro consideramos la siguiente manera de clasificar dicha investigación: por su dimensión temporal o el número de momentos o puntos en el tiempo en los cuales se recolectan datos.

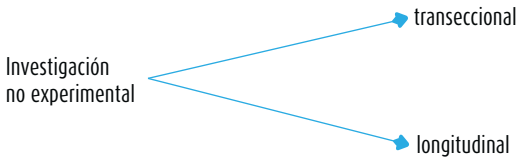
En algunas ocasiones la investigación se centra en:

a) Analizar cuál es el nivel o modalidad de una o diversas variables en un momento dado.
b) Evaluar una situación, comunidad, evento, fenómeno o contexto en un punto del tiempo.
c) Determinar o ubicar cuál es la relación entre un conjunto de variables en un momento.

En estos casos el diseño apropiado (con un enfoque no experimental) es el transversal o transeccional. Ya sea que su alcance inicial o final sea exploratorio, descriptivo, correlacional o explicativo.

Otras veces, la investigación se enfoca en: *a*) estudiar cómo evolucionan una o más variables o las relaciones entre ellas, o *b*) analizar los cambios al paso del tiempo de un evento, comunidad, proceso, fenómeno o contexto. En situaciones como éstas el diseño apropiado (en un enfoque no experimental) es el longitudinal.

Dicho de otro modo, los **diseños no experimentales** se pueden clasificar en transeccionales y longitudinales.



Investigación transeccional o transversal

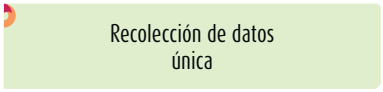
Diseños transeccionales (transversales) Investigaciones que recopilan datos en un momento único.

Los diseños de **investigación transeccional** o **transversal** recolectan datos en un solo momento, en un tiempo único (Liu, 2008 y Tucker, 2004). Su propósito es describir variables y analizar su incidencia e interrelación en un momento dado. Es como "tomar una fotografía" de algo que sucede. Por ejemplo:

1. Medir las percepciones y actitudes de mujeres jóvenes (18-25 años) que fueron abusadas sexualmente en el último mes en una urbe latinoamericana.
2. Evaluar el estado de los edificios de un barrio o una colonia, después de un terremoto.
3. Analizar el efecto que sobre la estabilidad emocional de un grupo de personas provocó en lo inmediato un acto terrorista (como el atentado del maratón de Boston en abril de 2013).
4. El estudio de Lee y Guerin (2009) para identificar si la satisfacción respecto a la calidad del diseño ambiental del interior de áreas de trabajo u oficinas afecta significativamente la satisfacción general del espacio de trabajo por parte de sus ocupantes y su desempeño laboral, en un momento específico.

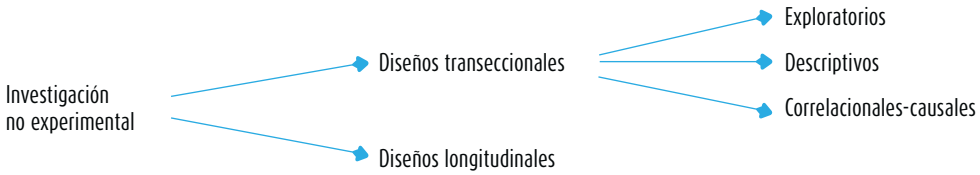
www.elosopanda.com | jamespoetrodriguez.com

Estos diseños se esquematizan de la siguiente manera:



Pueden abarcar varios grupos o subgrupos de personas, objetos o indicadores; así como diferentes comunidades, situaciones o eventos. Por ejemplo, analizar el efecto que sobre la estabilidad emocional provocó dicho acto terrorista en niños, adolescentes y adultos. Pero siempre, la recolección de los datos ocurre en un momento único.

A su vez, los **diseños transeccionales** se dividen en tres: exploratorios, descriptivos y correlacionales-causales.



Diseños transeccionales exploratorios

El propósito de los **diseños transeccionales exploratorios** es comenzar a conocer una variable o un conjunto de variables, una comunidad, un contexto, un evento, una situación. Se trata de una exploración inicial en un momento específico. Por lo general, se aplican a problemas de investigación nuevos o poco conocidos; además, constituyen el preámbulo de otros diseños (no experimentales y experimentales).

Por ejemplo, unas investigadoras pretenden obtener un panorama sobre el grado en que las empresas de una ciudad contratan a personas con capacidades distintas (impedimentos físicos, deficiencias motrices, visuales, mentales). Buscan en los archivos municipales y encuentran muy poca información, acuden a las cámaras empresariales de la localidad y tampoco descubren datos que les sean útiles. Entonces inician un sondeo en las organizaciones productivas de su localidad, haciendo una serie de preguntas a los gerentes de personal, recursos humanos o equivalentes: ¿contratan a personas con capacidades diferentes? ¿Cuántas personas al año, al mes? ¿Para qué tipo de empleos?, etc. Al explorar la situación logran formarse una idea del problema que les interesa y sus resultados son exclusivamente válidos para el tiempo y lugar en que efectuaron su estudio. Sólo recolectaron datos una vez. Posteriormente podrían planear una investigación descriptiva más profunda sobre la base proporcionada por esta primera aproximación o comenzar un estudio que indague qué empresas son las que contratan a más individuos con capacidades distintas y por qué motivos.

Diseños transeccionales descriptivos

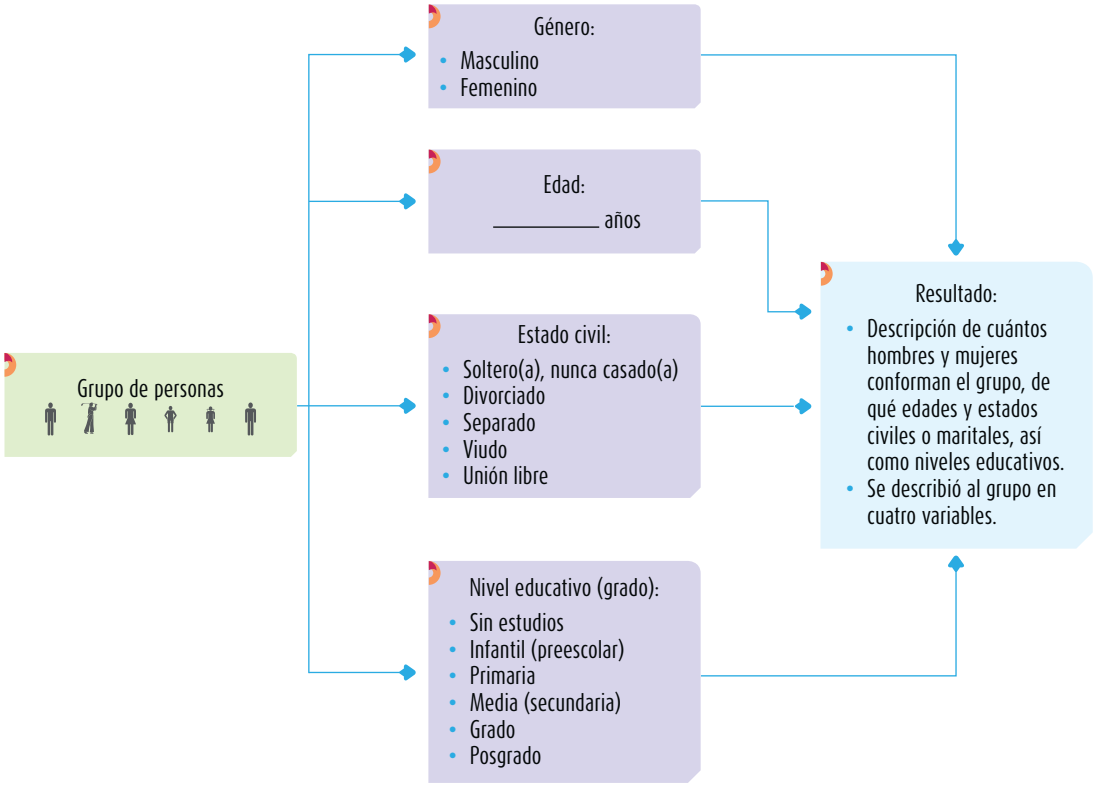
Los **diseños transeccionales descriptivos** tienen como objetivo indagar la incidencia de las modalidades o niveles de una o más variables en una población. El procedimiento consiste en ubicar en una o diversas variables a un grupo de personas u otros seres vivos, objetos, situaciones, contextos, fenómenos, comunidades, etc., y proporcionar su descripción. Son, por tanto, estudios puramente descriptivos y cuando establecen hipótesis, éstas son también descriptivas (de pronóstico de una cifra o valores).

Diseños transeccionales descriptivos Indagan la incidencia de las modalidades, categorías o niveles de una o más variables en una población, son estudios puramente descriptivos.

Por ejemplo: ubicar a un grupo de personas en las variables género, edad, estado civil o marital y nivel educativo.[15] Esto podría representarse así:

[15] El nivel educativo varía entre diferentes países, en algunos casos la educación media se refiere a secundaria y preparatoria; en otros, a secundaria o únicamente bachillerato.

Figura 7.6 Ejemplo de ubicación de personas.



En ciertas ocasiones, el investigador pretende realizar descripciones comparativas entre grupos o subgrupos de personas u otros seres vivos, objetos, comunidades o indicadores (esto es, en más de un grupo). Por ejemplo, un investigador que deseara describir el nivel de empleo en tres ciudades (Valencia, Caracas y Trujillo, en Venezuela).

Ejemplos

1. Las famosas encuestas nacionales de opinión sobre las tendencias de los votantes durante periodos electorales. Su objetivo es describir —en una elección específica— el número de votantes que se inclinan por los diferentes candidatos contendientes. Es decir, se centran en la descripción de las preferencias del electorado.
2. Un análisis sobre la tendencia ideológica de los 15 diarios de mayor tirada en América Latina. El foco de atención es únicamente describir, en un momento dado, cuál es la tendencia ideológica (de izquierda o derecha) de dichos periódicos. No se tiene como objetivo ver por qué manifiestan una u otra ideología, sino tan sólo describirlas.
3. Una investigación para evaluar los niveles de satisfacción de los clientes de un hotel respecto al servicio que reciben (no busca evaluar si las mujeres están más satisfechas que los hombres, ni asociar el nivel de satisfacción con la edad o los ingresos de los clientes).

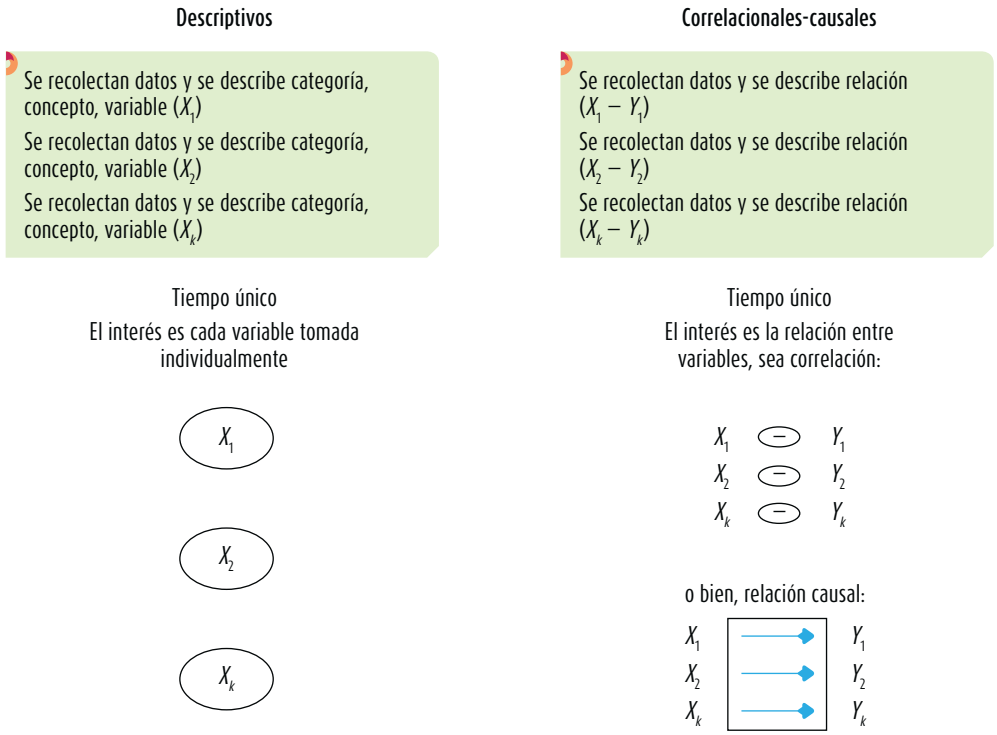
Imagine que su único propósito es describir físicamente a una persona (digamos, a Alexis, un joven de 14 años). Entonces, nos diría cuál es su estatura, talla, de qué color es su cabello y ojos, cómo es su complexión, etc. Así son los estudios descriptivos y queda claro que ni siquiera cabe la noción de manipulación, puesto que cada variable o concepto se trata individualmente: no se vinculan variables. Además, la descripción de Alexis es a la edad de 14 años (un solo momento), la cual variará en diferentes cuestiones conforme crezca (estatura, por ejemplo).

Diseños transeccionales correlacionales-causales

Estos diseños describen relaciones entre dos o más categorías, conceptos o variables en un momento determinado. A veces, únicamente en términos correlacionales, otras en función de la relación causa-efecto (causales).

La diferencia entre los diseños transeccionales descriptivos y los **diseños correlacionales-causales** se expresa gráficamente en la figura 7.7.

Figura 7.7 Comparación de diseños transeccionales descriptivos y correlacionales-causales.



Por tanto, los **diseños correlacionales-causales** pueden limitarse a establecer relaciones entre variables sin precisar sentido de causalidad o pretender analizar relaciones causales. Cuando se limitan a relaciones no causales, se fundamentan en planteamientos e hipótesis correlacionales; del mismo modo, cuando buscan evaluar vinculaciones causales, se basan en planteamientos e hipótesis causales. Veamos algunos ejemplos.

Ejemplo

1. Una investigación que pretendiera indagar la relación entre la atracción y la confianza durante el noviazgo en parejas de jóvenes, observando cuán vinculadas están ambas variables (se limita a ser correlacional).
2. Una investigación que estudiara cómo la motivación intrínseca influye en la productividad de los trabajadores de línea de grandes empresas industriales, de determinado país y en cierto momento, observando si los obreros más productivos son los más motivados; en caso de que así sea, evaluando por qué y cómo es que la motivación intrínseca contribuye a incrementar la productividad (esta investigación establece primero la correlación y luego la relación causal entre las variables).
3. Un estudio sobre la relación entre urbanización y alfabetismo en una nación latinoamericana, para ver qué variables macrosociales median en tal relación (causal).
4. Un estudio que pretendiera analizar quiénes compran más en las tiendas de una cadena departamental, los hombres o las mujeres, y de qué edades y perfiles socioeconómicos (correlacional: asocia nivel de compra con género, edad y nivel socioeconómico).

De los ejemplos se desprende lo que se ha comentado anteriormente: que en ciertas ocasiones sólo se pretende correlacionar categorías, variables, objetos o conceptos; pero en otras, se busca establecer relaciones causales. Debemos recordar que la causalidad implica correlación, pero no toda correlación significa causalidad.

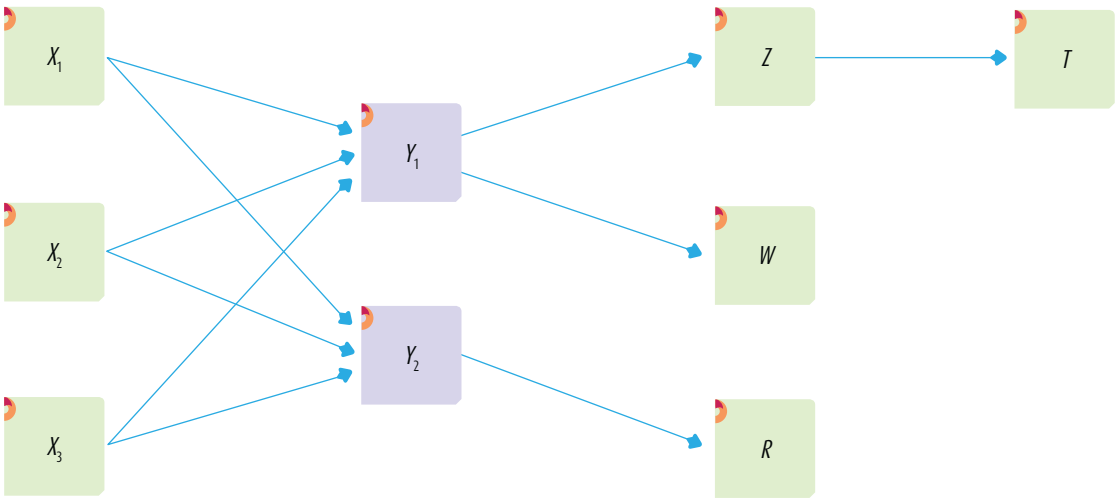
Estos diseños pueden ser sumamente complejos y abarcar diversas categorías, conceptos o variables. Cuando establecen relaciones causales son explicativos. Su diferencia con los experimentos es la base de la distinción entre experimentación y no experimentación. En los **diseños transeccionales correlacionales-causales**, las causas y los efectos ya ocurrieron en la realidad (estaban dados y manifestados) o suceden durante el desarrollo del estudio, y quien investiga los observa y reporta. En cambio, en los diseños experimentales y cuasiexperimentales se provoca intencionalmente al menos una causa y se analizan sus efectos o consecuencias.

Diseños transeccionales correlacionales-causales Describen relaciones entre dos o más categorías, conceptos o variables en un momento determinado, ya sea en términos correlacionales, o en función de la relación causa-efecto.

En todo estudio, la posible causalidad la establece el investigador de acuerdo con sus hipótesis, las cuales se fundamentan en la revisión de la literatura. En los experimentos —como ya se ha insistido—, la causalidad va en el sentido del tratamiento o tratamientos (variable o variables independientes) hacia el efecto o efectos (variable o variables dependientes). En los estudios transeccionales correlacionales-causales, la causalidad ya existe, pero es el investigador quien determina su dirección y establece cuál es la causa y cuál el efecto (o causas y efectos). Ya sabemos que para establecer un nexo causal: *a*) la o las variables independientes deben anteceder en tiempo a la o las dependientes, aunque sea por milésimas de segundo (por ejemplo, en la relación entre "el nivel de estudio de los padres" y "el interés por la lectura de los hijos", es obvio que la primera variable antecede a la segunda); y *b*) debe existir covariación entre la o las variables independientes y dependientes; pero además: *c*) la causalidad tiene que ser verosímil (si decidimos que existe un vínculo causal entre las variables "nutrición" y "rendimiento escolar", resulta lógico que la primera es causa de la segunda, pero no a la inversa).

Un **diseño correlacional-causal** puede limitarse a dos categorías, conceptos o variables, o incluso abarcar modelos o estructuras tan complejas como lo muestra la figura 7.8 (donde cada letra en recuadro representa una variable, un concepto, etcétera).

Figura 7.8 Ejemplo de una estructura de un diseño correlacional-causal complejo.



Asimismo, en ocasiones los diseños correlacionales-causales describen relaciones en uno o más grupos o subgrupos y suelen describir primero las variables incluidas en la investigación, para luego establecer las relaciones entre éstas (en primer lugar, son descriptivos de variables individuales, pero luego van más allá de las descripciones y establecen relaciones).

www.elosopanda.com | jamespoetrodriguez.com

Ejemplo

En una investigación para evaluar la credibilidad de tres conductores de televisión, y relacionar esta variable con el género, la ocupación y el nivel socioeconómico del teleauditorio. Primero, mediríamos qué tan creíble es cada conductor y describiríamos la credibilidad de los tres. Determinaríamos el género de las personas e investigaríamos su ocupación y nivel socioeconómico. Así, describiríamos estos tres elementos del teleauditorio. Posteriormente, relacionaríamos la credibilidad y el género (para ver si hay diferencias entre hombres y mujeres en cuanto a la credibilidad de los tres conductores), la credibilidad y la ocupación (con el fin de analizar si los conductores tienen una credibilidad similar o diferente entre las distintas ocupaciones) y la credibilidad y el nivel socioeconómico (para evaluar diferencias por este nivel). De este modo, primero describimos y luego correlacionamos.

En estos diseños, en su modalidad únicamente causal, a veces se reconstruyen las relaciones a partir de las variables dependientes, en otras a partir de las independientes y en otras más sobre la base de variabilidad amplia de las independientes y dependientes (León y Montero, 2003). Al primer caso se les conoce como diseños *retrospectivos,* al segundo como *prospectivos* y al tercero como de *causalidad múltiple*. En el capítulo 5 adicional, "Diseños: segunda parte", que puede descargarse del centro de recursos en línea, hay ejemplos de estos diseños.

Centro de recursos en línea

Encuestas de opinión

Las **encuestas de opinión** son consideradas por diversos autores como un diseño o método.[16] En la clasificación de la presente obra serían consideradas investigaciones no experimentales transversales o transeccionales descriptivas o correlacionales-causales, ya que a veces tienen los propósitos de unos u otros diseños y a veces de ambos (Archester, 2005). Generalmente utilizan cuestionarios que se aplican en diferentes contextos (entrevistas en persona, por medios electrónicos como correos o páginas web, en grupo, etc.). El proceso de una **encuesta de opinión** y ejemplos de ella se comentan en el capítulo 6 adicional: "Encuestas (*surveys)*", del centro de recursos en línea.

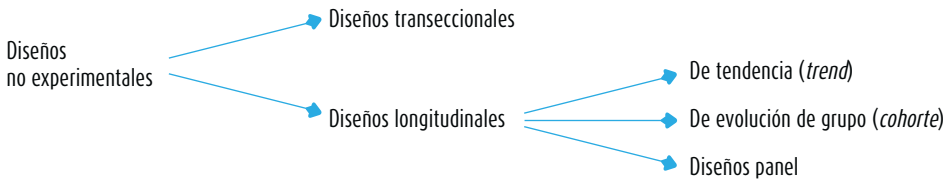
Centro de recursos en línea

Investigación longitudinal o evolutiva

En ocasiones, el interés del investigador es analizar cambios al paso del tiempo en determinadas categorías, conceptos, sucesos, variables, contextos o comunidades, o bien, de las relaciones entre éstas. Aún más, a veces ambos tipos de cambios. Entonces disponemos de los **diseños longitudinales**, los cuales recolectan datos en diferentes momentos o periodos para hacer inferencias respecto al cambio, sus determinantes y consecuencias. Tales puntos o periodos generalmente se especifican de antemano. Por ejemplo, un investigador que buscara analizar cómo evolucionan los niveles de empleo durante cinco años en una ciudad; otro que pretendiera estudiar cómo ha cambiado el contenido sexual en las telenovelas de cierto país en los últimos 10 años, y uno más que buscara observar cómo se desarrolla una comunidad indígena con los años, con la llegada de la computadora e internet a su vida. Son, pues, estudios de seguimiento.

Diseños longitudinales Estudios que recaban datos en diferentes puntos del tiempo, para realizar inferencias acerca de la evolución del problema de investigación o fenómeno, sus causas y sus efectos.

Los **diseños longitudinales** se dividen en tres tipos: *diseños de tendencia*, *diseños de análisis evolutivo de grupos* (*cohorte*) y *diseños panel*, como se indica en el siguiente esquema:

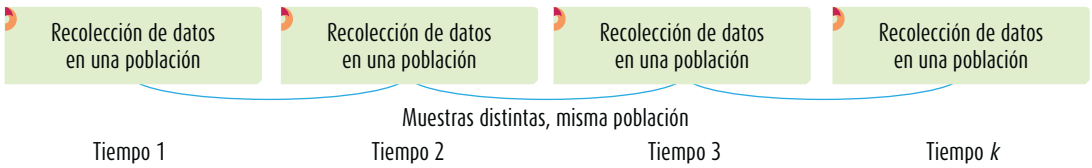


[16] Por ejemplo: McLaren (2014); Hernández-Sampieri *et al.* (2013); Creswell (2013a); Mertens (2010); Roberts (2009); Groves *et al.* (2009); Bowers (2008) y Julien (2008).

Diseños longitudinales de tendencia

Los **diseños de tendencia** son aquellos que analizan cambios al paso del tiempo en categorías, conceptos, variables o sus relaciones de alguna población en general. Su característica distintiva es que la atención se centra en la población o universo. Por ejemplo, una investigación para analizar cambios en la actitud hacia el aborto por parte de universitarios de una comunidad. Dicha actitud se mide en varios puntos en el tiempo (digamos, cada año o en periodos no preestablecidos durante 10 años) y se examina su evolución a lo largo de este gran periodo. Se puede estudiar a toda la población, o bien tomar una muestra de ella cada vez que se observen o midan las variables o las relaciones entre éstas. Es importante señalar que los participantes o casos de la investigación **no** son los mismos, pero la población **sí**. Obviamente, los universitarios crecen con el tiempo, pero siempre hay una población de ellos. Por ejemplo, los estudiantes de Medicina de la Universidad Complutense de Madrid de hoy no serán las mismas personas que las de años futuros, pero siempre habrá una población de estudiantes de medicina de dicha institución. El principio se aplica también a seres que experimentan crecimiento, como los animales o las plantas. Estos diseños se representan en la figura 7.9.

Figura 7.9 Esquema de un diseño longitudinal de tendencia.



Ejemplo

Analizar la manera en que evoluciona la percepción sobre tener relaciones sexuales premaritales en las mujeres jóvenes adultas (20 a 25 años) de Valledupar, Colombia, de aquí al año 2024. Las mujeres aumentan su edad, pero siempre habrá una población de mujeres de esas edades en tal ciudad. Las participantes seleccionadas son otras, pero el universo es el mismo.

Diseños longitudinales de evolución de grupo (*cohortes*)

Diseños de tendencia y de evolución de grupo Ambas clases de diseños monitorean cambios en una población o subpoblación a través del tiempo, usando una serie de muestras que abarcan a diferentes participantes en cada ocasión, pero en los primeros la población es la misma y en los segundos se toma como universo a los sobrevivientes de la población.

Con los **diseños de evolución de grupo** se examinan cambios a través del tiempo en subpoblaciones o grupos específicos. Su atención son las *cohortes* o grupos de individuos vinculados de alguna manera o identificados por una característica común, generalmente la edad o la época o la región geográfica (Bell, 2009; Hsieh, 2007 y Glenn, 1977). Un ejemplo de estos grupos (*cohortes*) sería el formado por las personas que nacieron en 1973 en Chile, el año del derrocamiento del gobierno de Salvador Allende; pero también podría utilizarse otro criterio de agrupamiento temporal, como las personas que se casaron durante 2010 en Rosario, Argentina; o los niños de la Ciudad de México que iban en primaria cuando ocurrió el gran terremoto de 1985. Tales diseños hacen seguimiento de los grupos al paso del tiempo y por lo común se extrae una muestra cada vez que se recolectan datos sobre el grupo o la subpoblación, más que incluir a toda la subpoblación.

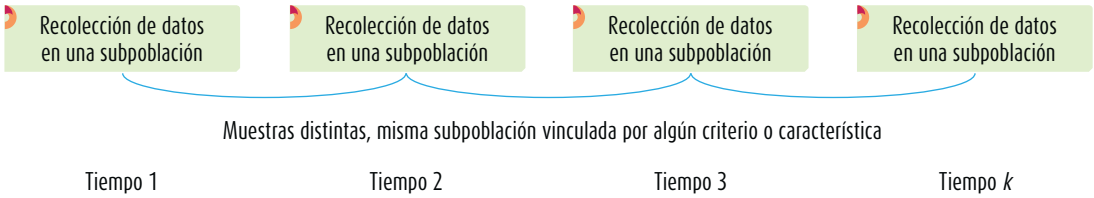
Ejemplo

Una investigación nacional sobre las actitudes hacia la democracia de los mexicanos nacidos en 2000 (recordemos que en México hasta dicho año no hubo elecciones presidenciales verdaderamente democráticas), digamos cada cinco años, comenzando a partir de 2020. En este año se obtendría una muestra de mexicanos de 20 años de edad y se medirían las actitudes. En 2025, se seleccionaría una muestra de mexicanos de 25 años y se medirían las actitudes. En 2030, se elegiría una muestra de mexicanos de 30 años, y así sucesivamente. De esta forma, se analizan la evolución y los cambios de las actitudes mencionadas. Desde luego que, aunque el conjunto específico de personas estudiadas en cada tiempo o medición llega a ser diferente, cada muestra representa a los sobrevivientes del grupo de mexicanos nacidos en 2000.

www.elosopanda.com | jamespoetrodriguez.com

Los diseños de evolución de grupo se pueden esquematizar como en la figura 7.10.

Figura 7.10 Esquema de los diseños de evolución de grupo.



Diseños longitudinales panel

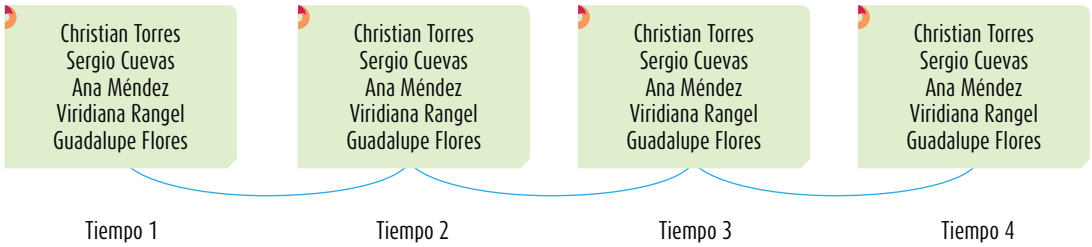
Los **diseños panel** son similares a las dos clases de diseños vistas anteriormente, sólo que los mismos casos o participantes son medidos u observados en todos los tiempos o momentos.[17]

Diseños panel Toda una población o grupo es seguido a través del tiempo.

Un ejemplo sería una investigación que observara anualmente los cambios en las actitudes (mediante la aplicación de una prueba estandarizada) de un grupo de ejecutivos en relación con un programa para elevar la productividad, por ejemplo, durante cinco años. Cada año se observaría la actitud de los mismos ejecutivos. Es decir, los individuos, y no sólo la muestra, población o subpoblación, son los mismos.

Otro ejemplo sería observar mensualmente (durante dos años) a un grupo que acude a psicoterapia para analizar si se incrementan sus expresiones verbales de exploración y discusión de planes futuros, y si disminuyen las de hechos pasados (en cada observación, los pacientes serían las mismas personas). La forma gráfica de representar este ejemplo de diseño longitudinal se muestra en la figura 7.11.

Figura 7.11 Ejemplo de diseño longitudinal panel.



Un ejemplo más de diseño panel consiste en analizar la evolución de pacientes de un determinado tipo de cáncer (de mama, pongamos como caso), en la que se evalúe al grupo durante cuatro etapas: la primera, un mes después de iniciar el tratamiento médico; la segunda, tres meses más tarde; la tercera, seis meses después del tratamiento, y la cuarta, al cabo de un año. Siempre se incluirán a las mismas pacientes con nombre y apellido, descartando a quienes fallecen.

Estos diseños se utilizan frecuentemente en el seguimiento de la conducta animal, colocando a ciertos ejemplares (leones, cobras, osos polares, etc.) dispositivos de rastreo.

En los diseños panel se tiene la ventaja de que, además de conocer los cambios grupales, se conocen los cambios individuales. Se sabe qué casos específicos introducen el cambio. La desventaja es que a veces resulta muy difícil obtener con exactitud a los mismos participantes para una segunda medición u observaciones subsecuentes. Este tipo de diseños sirve para estudiar poblaciones o grupos más específicos y es conveniente cuando se tienen poblaciones relativamente estáticas.

[17] Andreß, Golsch y Schmidt (2013); Baltagi (2008), Keulenaer (2008); Kim (2008); Feng, Tang y Tu (2007) y Tucker (2004).

Por otra parte, deben verse con cuidado los efectos que una medición, un registro o una observación llega a tener sobre otras posteriores (recuérdese el efecto de administración de la prueba vista como fuente de invalidación interna en experimentos y cuasiexperimentos, sólo que aplicada al contexto no experimental). Los diseños panel podrían esquematizarse como se puede observar en la figura 7.12.

Figura 7.12 Esquema de diseño panel.

Los **diseños longitudinales** se fundamentan en hipótesis de diferencia de grupos, correlacionales y causales. Estos diseños recolectan datos sobre categorías, sucesos, comunidades, contextos, variables o sus relaciones, en dos o más momentos, para evaluar el cambio en éstas. Ya sea al tomar a una población (diseños de tendencia), a una subpoblación (diseños de análisis evolutivo de un grupo o *cohorte*) o a los mismos casos o participantes (diseños panel). Ejemplos de temas serían: resistencia de materiales para construir edificios a través del tiempo, recaudación fiscal en distintos años, comportamiento de acciones en la bolsa de valores de una nación antes y después de algunos sucesos, duración de algún material para cubrir picaduras o daños a los molares, la relación entre el clima y la cultura organizacionales durante varios periodos, o los impactos después de una guerra (a mediano y largo plazos) en alguna sociedad del siglo XVI (histórico).

Comparación de los diseños transeccionales y longitudinales

Los estudios longitudinales tienen la ventaja de que proporcionan información sobre cómo las categorías, conceptos, procesos, variables, comunidades, fenómenos, y sus relaciones evolucionan al paso del tiempo. Sin embargo, suelen ser más costosos que los transeccionales. La elección de un tipo de diseño u otro depende más bien de los propósitos de la investigación y de su alcance.

¿Cuáles son las características de la investigación no experimental en comparación con la investigación experimental?

Una vez más recalcamos que tanto la **investigación experimental** como la **no experimental** son herramientas muy valiosas y ningún tipo es mejor que el otro. El diseño que se elige en una investigación depende más bien del problema que se quiere resolver y del contexto del estudio. Desde luego, ambos tipos de investigación poseen características propias que es necesario resaltar.

El control sobre las variables es más riguroso en los experimentos que en los diseños cuasiexperimentales y, a su vez, estas dos clases de estudios logran mayor control que los diseños no experimentales. En un experimento se analizan relaciones "puras" entre las variables de interés, sin contaminación de otras variables y, por ello, es posible establecer relaciones causales con mayor precisión. Por ejemplo, en un experimento sobre el aprendizaje variaríamos el estilo de liderazgo del profesor, el método de enseñanza y otros factores. Así, sabríamos cuánto afectó cada variable. En cambio, en la investigación no experimental resulta más complejo separar los efectos de las múltiples variables que intervienen, sin embargo, puede hacerse por inferencia.

Por lo que respecta a la posibilidad de repetición, prácticamente todos los diseños pueden replicarse, aunque en los longitudinales es mucho más complejo.

www.elosopanda.com | jamespoetrodriguez.com

Ahora bien, como menciona Kerlinger (1979), en **los experimentos** (sobre todo en los de laboratorio) las variables independientes pocas veces tienen tanta fuerza como en la realidad o la cotidianidad. Es decir, en el laboratorio tales variables no muestran la verdadera magnitud de sus efectos, la cual suele ser mayor fuera del laboratorio. Por tanto, si se encuentra un efecto en el laboratorio, éste tenderá a ser mayor en la realidad.

En cambio, en la **investigación no experimental** estamos más cerca de las variables formuladas hipotéticamente como "reales" y, en consecuencia, tenemos mayor validez externa (posibilidad de generalizar los resultados a otros individuos y situaciones comunes).

Investigación experimental y no experimental Se utilizan para el avance del conocimiento y en ocasiones resulta más apropiado un tipo u otro, dependiendo del problema de investigación al que nos enfrentemos.

En el caso de los experimentos donde se involucran seres humanos, una desventaja es que normalmente se selecciona un número de individuos poco o medianamente representativo respecto a las poblaciones que se estudian. La mayoría de los experimentos utiliza muestras no mayores de 200 personas, lo que dificulta la generalización de resultados a poblaciones más amplias. Por tal razón, los resultados de un experimento deben observarse con precaución y es por medio de la réplica de éste (en distintos contextos y con diferentes individuos) como van generalizándose dichos resultados. Cuando se experimenta con objetos (por ejemplo materiales), como son menos heterogéneos, se requieren muestras menores. Con el fin de vincular los alcances del estudio, las hipótesis y el diseño, sugerimos se considere la tabla 7.5.

Diversos problemas de investigación se pueden abordar experimental y no experimentalmente. Por ejemplo, si deseáramos analizar la relación entre la motivación y la productividad en los trabajadores de cierta empresa, seleccionaríamos un conjunto de éstos y lo dividiríamos al azar en cuatro grupos: uno en el que se propicie una elevada motivación, otro con mediana motivación, otro más con baja motivación y un último al que no se le administre ningún motivador. Después compararíamos la productividad de los grupos. Tendríamos un experimento.

Tabla 7.5 Correspondencia entre tipos de estudio, hipótesis y diseño de investigación

Estudio	Hipótesis	Posibles diseños
Exploratorio	• No se establecen, lo que se puede formular son conjeturas iniciales	• Transeccional exploratorio o descriptivo • Preexperimental
Descriptivo	• Descriptiva	• Preexperimental • Transeccional descriptivo
Correlacional	• Diferencia de grupos sin atribuir causalidad	• Cuasiexperimental • Transeccional correlacional • Longitudinal (no experimental)
	• Correlacional	• Cuasiexperimental • Transeccional correlacional • Longitudinal (no experimental)
Explicativo	• Diferencia de grupos atribuyendo causalidad	• Experimental puro • Cuasiexperimental, longitudinal y transeccional causal (cuando hay bases para inferir causalidad, un mínimo de control y análisis estadísticos apropiados para relaciones causales)
	• Causales	• Experimental puro • Cuasiexperimental, longitudinal y transeccional causal (cuando hay bases para inferir causalidad, un mínimo de control y análisis estadísticos apropiados para relaciones causales)

Si se tratara de grupos intactos (turnos) tendríamos un cuasiexperimento. En cambio, si midiéramos la motivación existente en los trabajadores, así como su productividad y relacionáramos ambas variables, estaríamos realizando una investigación transeccional correlacional. Y si cada seis meses midiéramos las dos variables y estableciéramos su correlación efectuaríamos un estudio longitudinal.

Los estudios de caso

Los **estudios de caso** son considerados por algunos autores como una clase de diseños, a la par de los experimentales, no experimentales y cualitativos (Creswell, 2013a; Hancock y Algozzine, 2011; Mertens, 2010; Aaltio y Heilmann, 2009 y Williams, Grinnell y Unrau, 2005), mientras que otros los ubican como una clase de diseño experimental (León y Montero, 2003) o un diseño etnográfico (Creswell, 2013b). También han sido concebidos como un asunto de muestreo o un método (Runeson, Host, Rainer y Regnell, 2012).

La realidad es que los **estudios de caso** son todo lo anterior (Yin, 2013; Blatter, 2008; Hammersley, 2003). Poseen sus propios procedimientos y clases de diseños. Los podríamos definir como "estudios que al utilizar los procesos de investigación cuantitativa, cualitativa o mixta analizan profundamente una unidad holística para responder al planteamiento del problema, probar hipótesis y desarrollar alguna teoría" (Hernández-Sampieri y Mendoza, 2008). Esta definición los sitúa más allá de un tipo de diseño o muestra, pero ciertamente es la más cercana a la evolución que han tenido los estudios de caso en los últimos años.

En ocasiones, **los estudios de caso** utilizan la experimentación, es decir, se constituyen en estudios preexperimentales. Otras veces se fundamentan en un diseño no experimental (transversal o longitudinal) y en ciertas situaciones se convierten en estudios cualitativos, al emplear métodos cualitativos. Asimismo, pueden valerse de las diferentes herramientas de la investigación mixta.

Centro de recursos en línea

Tales estudios en sus principales modalidades son comentados en el capítulo 4 adicional que el lector podrá descargar del centro de recursos en línea en *Material complementario* → *Capítulos* → *Estudios de caso*, y que dada su importancia merecen una atención particular.

Por ahora mencionaremos que la unidad o caso investigado puede tratarse de un individuo, una pareja, una familia, un objeto (una pirámide como la de Keops, un material radiactivo), un sistema (fiscal, educativo, terapéutico, de capacitación, de trabajo social), una organización (hospital, fábrica, escuela), un hecho histórico, un desastre natural, un proceso de manufactura, una comunidad, un municipio, un departamento o estado, una nación, etc. En el capítulo 4 adicional, "Estudios de caso", incluso se trata un ejemplo de una investigación de una persona que padecía lupus eritematoso sistémico con 31 años de evolución, que mezcla aspectos experimentales con elementos cualitativos.

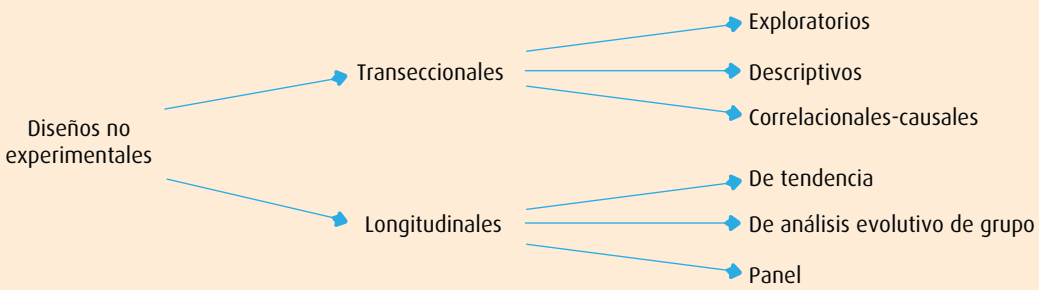
En la tabla 7.6 se encuentran preguntas de investigación que corresponderían a estudios de caso.

Tabla 7.6 Posibles estudios de caso derivados de preguntas de investigación

Preguntas de investigación
¿Qué funciones sociales o religiosas cumplía la construcción primitiva de Stonehenge en Sollysbury, Inglaterra? (Unidad o caso: un objeto o construcción).
¿Por qué se divorciaron Lupita y Adrián? (Unidad: pareja).
¿Cuáles fueron las causas que provocaron el desplome de un avión determinado? (Unidad: desastre aéreo).
¿Cuáles son las razones que llevaron a un estado de esquizofrenia a Carlos Codolla? (Unidad: individuo).
¿Quién sería el asesino en un determinado crimen? (Unidad: evento).
¿Cómo era la personalidad de Robert F. Kennedy? (Unidad: personaje histórico).
¿Qué daños a la infraestructura de cierta comunidad causó el gran Tsunami de 2004? (Unidad: evento o catástrofe).
¿Cómo puede caracterizarse el clima organizacional de la empresa Lucymex? (Unidad: organización).

Resumen

- El "diseño" se refiere al plan o la estrategia concebido para obtener la información que se desea.
- En el caso del proceso cuantitativo, el investigador utiliza su diseño para analizar la certeza de las hipótesis formuladas en un contexto específico o para aportar evidencia respecto de los lineamientos de la investigación (si es que no se tienen hipótesis).
- En un estudio pueden plantearse o tener cabida uno o más diseños.
- La tipología propuesta clasifica a los diseños en experimentales y no experimentales.
- Los diseños experimentales se dividen en experimentos "puros", cuasiexperimentos y preexperimentos.
- Los diseños no experimentales se dividen por el número de veces que recolectan datos en transeccionales y longitudinales.
- En su acepción más general, un experimento consiste en aplicar un estímulo, intervención o tratamiento a un caso, situación, fenómeno, individuo o grupo, y ver el efecto de ese estímulo en alguna o algunas variables. Esta observación se puede realizar en condiciones de mayor o menor control. El máximo control se alcanza en los experimentos "puros".
- Deducimos que un tratamiento afectó cuando observamos diferencias (en las variables que supuestamente serían las afectadas) entre un caso, grupo o fenómeno al que se le administró dicho estímulo y un caso, grupo o fenómeno al que no se le administró, siendo ambos iguales en todo, excepto en esto último.
- La variable independiente es la causa y la dependiente el efecto.
- Para lograr el control o la validez interna, los casos, fenómenos o grupos que se comparen deben ser iguales en todo, menos en el hecho de que a un caso, fenómeno o grupo se le aplicó el estímulo y al otro no. A veces graduamos la cantidad del estímulo que se administra, es decir, a distintos casos, fenómenos o grupos (semejantes) les aplicamos diferentes grados del estímulo para observar si provocan efectos distintos.
- La asignación al azar es normalmente el método preferible para lograr que los casos, fenómenos o grupos del experimento sean comparables (semejantes).
- Las principales fuentes que pueden invalidar un experimento son: historia, maduración, inestabilidad, administración de pruebas, instrumentación, regresión, selección, mortalidad experimental, difusión de tratamientos experimentales, compensación y el experimentador.
- Los experimentos que hacen equivalentes a los casos, fenómenos o grupos y que mantienen esta equivalencia durante el desarrollo de aquellos controlan las fuentes de invalidación interna.
- Lograr la validez interna es el objetivo metodológico y principal de todo experimento. Una vez que se consigue, es ideal alcanzar validez externa (posibilidad de generalizar los resultados a la población, otros experimentos y situaciones no experimentales).
- Las principales fuentes de invalidación externa son: efecto reactivo de las pruebas, efecto de interacción entre los errores de selección y el tratamiento experimental, efectos reactivos de los tratamientos experimentales, interferencia de tratamientos múltiples, imposibilidad de replicar los tratamientos, descripciones insuficientes del tratamiento experimental, efectos de novedad e interrupción, el experimentador, interacción entre la historia o el lugar y los efectos del tratamiento experimental, mediciones de la variable dependiente.
- Hay dos contextos en los que se realizan los experimentos: el laboratorio y el campo.
- En los cuasiexperimentos no se asignan al azar los sujetos a los grupos experimentales, sino que se trabaja con grupos intactos.
- Los cuasiexperimentos alcanzan validez interna en la medida en que demuestran la equivalencia inicial de los casos, fenómenos o grupos participantes y la equivalencia en el proceso de experimentación.
- Los experimentos "puros" constituyen estudios explicativos; los preexperimentos básicamente son estudios exploratorios y descriptivos; los cuasiexperimentos son, fundamentalmente, correlacionales aunque pueden llegar a ser explicativos.
- La investigación no experimental es la que se realiza sin manipular deliberadamente las variables independientes; se basa en categorías, conceptos, variables, sucesos, fenómenos o contextos que ya ocurrieron o se dieron sin la intervención directa del investigador.
- La investigación no experimental también se conoce como investigación *ex post-facto* (los hechos y variables ya ocurrieron), y observa variables y relaciones entre éstas en su contexto natural.
- Los diseños no experimentales se dividen de la siguiente manera:



- Los diseños transeccionales realizan observaciones en un momento o tiempo único. Cuando recolectan datos sobre una nueva área sin ideas prefijadas y con apertura son más bien exploratorios; cuando recolectan datos sobre cada una de las categorías, conceptos, variables, contextos, comunidades o fenómenos, e informan lo que arrojan esos datos son descriptivos; cuando además describen vinculaciones entre categorías, conceptos, variables, sucesos, contextos o fenómenos son correlacionales, y si establecen procesos de causalidad entre tales términos se consideran correlacionales-causales (explicativos).
- Las encuestas de opinión son investigaciones no experimentales transversales o transeccionales descriptivas o correlacionales-causales, ya que a veces tienen los propósitos de unos u otros diseños y a veces de ambos.
- En los diseños transeccionales, en su modalidad "causal", a veces se reconstruyen las relaciones a partir de las variables dependientes, en otras a partir de las independientes y en otras más sobre la base de variabilidad amplia de las independientes y dependientes (al primer caso se les conoce como "retrospectivos", al segundo como "prospectivos" y al tercero como "causalidad múltiple"). Este tema se revisa en el capítulo 5 adicional del centro de recursos en línea, Diseños: segunda parte.
- Los diseños longitudinales sirven para efectuar observaciones en dos o más momentos o puntos en el tiempo. Si estudian una población son diseños de tendencia, si analizan una subpoblación o grupo específico son diseños de análisis evolutivo de grupo (*cohorte*) y si se estudian los mismos casos o participantes son diseños panel.
- El tipo de diseño a elegir se encuentra condicionado por el enfoque seleccionado, el problema a investigar, el contexto que rodea la investigación, los alcances del estudio y las hipótesis formuladas.

Conceptos básicos

- Alcances del estudio y diseño
- Casos
- Cohorte
- Control experimental
- Cuasiexperimento
- Diseño
- Diseño experimental
- Diseño no experimental
- Diseños longitudinales
- Diseños transeccionales
- Estímulo, intervención o tratamiento experimental/ manipulación de la variable independiente
- Experimento
- Experimento de campo
- Experimento de laboratorio
- Fuentes de invalidación interna
- Fuentes de invalidación externa
- Grupos intactos
- Investigación *ex post-facto*
- Panel
- Participantes (sujetos) del experimento
- Preexperimento
- Población (universo)
- Subpoblación
- Validez externa
- Validez interna
- Variable dependiente
- Variable experimental
- Variable independiente

Ejercicios

1. Seleccione una serie de variables y piense cómo se manipularían en situaciones experimentales. ¿Cuántos niveles o modalidades podrían incluirse para cada variable? ¿Estos niveles o modalidades cómo podrían traducirse en tratamientos experimentales? ¿Se tendría un nivel de ausencia (cero) de la variable independiente? ¿En qué consistiría éste?
2. Centro de recursos en línea Seleccione un experimento en alguna revista académica (busque en el centro de recursos en línea: Material complementario → Apéndices → Apéndice 1: "Publicaciones periódicas más importantes"). Analice: ¿cuál es el planteamiento del problema (objetivos y preguntas de investigación)? ¿Cuál es la hipótesis que se busca probar por medio de los resultados del experimento? ¿Cuál es la variable independiente o cuáles son las variables independientes? ¿Cuál es la variable o las variables dependientes? ¿Cuántos casos, fenómenos o grupos se incluyen en el experimento? ¿Son equivalentes? ¿Cuál es el diseño que el autor o autores han elegido? ¿Se controlan las fuentes de invalidación interna? ¿Se controlan las fuentes de invalidación externa? ¿Se encontró algún efecto?
3. Un grupo de investigadores intenta analizar el efecto que tiene la extensión de un discurso político sobre la actitud hacia el tema tratado y al orador. La extensión del discurso es la variable independiente y tiene cuatro niveles: 15 minutos, media hora, una hora y 90 minutos. Las variables dependientes son la actitud hacia el orador (favorable-desfavorable)

5.3.2. Selección de técnicas e instrumentos para la recolección de datos

Para el desarrollo del subtema **"Selección de técnicas e instrumentos para la recolección de datos"**, se consideran aquellos contenidos relacionados con la identificación, diferenciación, selección y utilización de las principales técnicas e instrumentos empleados para obtener información en una investigación. Se priorizan los contenidos que permiten comprender la **relación entre técnica e instrumento**, las principales técnicas de recolección de datos —como la **observación, la encuesta y la entrevista**—, así como el **cuestionario** y los criterios fundamentales para determinar la calidad de los instrumentos, especialmente su **confiabilidad, validez y objetividad**. De esta manera, los aportes de **Fidias G. Arias** y **Hernández Sampieri** se complementan para abordar el subtema de manera clara y suficiente, evitando incorporar contenidos correspondientes a otras tareas metodológicas.

N.°	Libros	Capítulo	Página	Contenido
1	*El Proyecto de Investigación* — **Fidias G. Arias**	**Capítulo 6. Técnicas e instrumentos de recolección de datos**	**p. 67**	**Relación entre técnica e instrumento:** diferenciación entre las técnicas utilizadas para obtener información y los instrumentos empleados para registrar o recopilar los datos.
2	*El Proyecto de Investigación* — **Fidias G. Arias**	**Capítulo 6. Técnicas e instrumentos de recolección de datos**	**p. 69**	**La observación y sus instrumentos:** utilización de la observación como técnica de recolección de datos y los instrumentos empleados para su registro.
3	*El Proyecto de Investigación* — **Fidias G. Arias**	**Capítulo 6. Técnicas e instrumentos de recolección de datos**	**p. 72**	**La encuesta y sus instrumentos:** características de la encuesta como técnica para obtener información y los instrumentos asociados a su aplicación.
4	*El Proyecto de Investigación* — **Fidias G. Arias**	**Capítulo 6. Técnicas e instrumentos de recolección de datos**	**p. 73**	**La entrevista y sus instrumentos:** características de la entrevista como técnica de recolección de información y los instrumentos utilizados para desarrollarla.
5	*El Proyecto de Investigación* — **Fidias G. Arias**	**Capítulo 6. Técnicas e instrumentos de recolección de datos**	**p. 74**	**El cuestionario:** características y utilización como instrumento para recopilar información mediante preguntas organizadas.
6	*Metodología de la Investigación* — **Hernández Sampieri, R., Fernández Collado, C. y Baptista Lucio, P.**	**Capítulo 9. Recolección de datos cuantitativos**	**p. 197**	**Requisitos que debe cubrir un instrumento de medición:** confiabilidad, validez y objetividad como características

				fundamentales para determinar la calidad y pertinencia del instrumento.
7	*Metodología de la Investigación* — **Hernández Sampieri, R., Fernández Collado, C. y Baptista Lucio, P.**	**Capítulo 9. Recolección de datos cuantitativos**	**p. 217**	**Tipos de instrumentos de medición o recolección de datos cuantitativos:** identificación y descripción de los principales instrumentos que pueden emplearse según las características y objetivos de la investigación.
8	*Metodología de la Investigación* — **Hernández Sampieri, R., Fernández Collado, C. y Baptista Lucio, P.**	**Capítulo 9. Recolección de datos cuantitativos**	**p. 217**	**El cuestionario como instrumento de recolección de datos:** características y utilización del cuestionario para recopilar información mediante preguntas estructuradas de acuerdo con los objetivos de la investigación.

Capítulo 6
Técnicas e Instrumentos de Recolección de Datos

"Un instrumento científico
abre una ventana al conocimiento."
Isaac Asimov

6.1. Relación entre técnica e instrumento

Una vez efectuada la operacionalización de las variables y definidos los indicadores, es hora de seleccionar las técnicas e instrumentos de recolección de datos pertinentes para verificar las hipótesis o responder las interrogantes formuladas. Todo en correspondencia con el problema, los objetivos y el diseño de investigación.

En este orden de ideas:

> Se entenderá por técnica de investigación, el procedimiento o forma particular de obtener datos o información.

Las técnicas son particulares y específicas de una disciplina, por lo que sirven de complemento al método científico, el cual posee una aplicabilidad general.

Diseño	Técnicas	Instrumentos	
Diseño de Investigación Documental	Análisis documental	Fichas Computadora y sus unidades de almacenaje	
	Análisis de contenido	Cuadro de registro y clasificación de las categorías	
Diseño de Investigación de Campo	Observación	Estructurada	Lista de cotejo Escala de estimación
		No Estructurada	Diario de campo Cámaras: fotográfica y de video
	Encuesta	Oral	Guía de encuesta (Tarjeta) Grabador Cámara de video
		Escrita	Cuestionario
	Entrevista	Estructurada	Guía de entrevista Grabador / Cámara de video
		No estructurada	Libreta de notas Grabador / Cámara de video

Ahora bien, la aplicación de una técnica conduce a la obtención de información, la cual debe ser guardada en un medio material de manera que los datos puedan ser recuperados, procesados, analizados e interpretados posteriormente. A dicho soporte se le denomina instrumento.

Un instrumento de recolección de datos es cualquier recurso, dispositivo o formato (en papel o digital), que se utiliza para obtener, registrar o almacenar información.

Son ejemplos de instrumentos:

- Un cuestionario en cuya estructura quedan registradas las respuestas suministradas por el encuestado.
- Una libreta en la que el investigador anota lo observado.
- Computadoras portátiles con sus respectivas unidades para almacenaje de información: disco duro, CD o memorias portátiles (pendrive).
- Dispositivos tales como cámara fotográfica y de video (filmadora), grabador de audio, etc.

6.2. La observación y sus instrumentos

La observación es una técnica que consiste en visualizar o captar mediante la vista, en forma sistemática, cualquier hecho, fenómeno o situación que se produzca en la naturaleza o en la sociedad, en función de unos objetivos de investigación preestablecidos.

Se hace especial referencia a la observación directa, ya que la indirecta se realiza a través de instrumentos muy sofisticados tales como: microscopio, telescopio, monitores, entre otros.

La observación puede ser:

a) Observación simple o no participante

Es la que se realiza cuando el investigador observa de manera neutral sin involucrarse en el medio o realidad en la que se realiza el estudio.

b) Observación participante

En este caso el investigador pasa a formar parte de la comunidad o medio donde se desarrolla el estudio.

Así mismo, la observación también se clasifica en:

- Observación libre o no estructurada

Es la que se ejecuta en función de un objetivo, pero sin una

guía prediseñada que especifique cada uno de los aspectos que deben ser observados.

- Observación estructurada

  Es aquella que además de realizarse en correspondencia con unos objetivos, utiliza una guía diseñada previamente, en la que se especifican los elementos que serán observados.

6.2.1. Instrumentos de la observación

En el caso de la observación libre o no estructurada, se emplean instrumentos tales como: diario de campo, libreta o cuaderno de notas, cámara fotográfica y cámara de video.

Para la observación estructurada, se utilizan instrumentos prediseñados tales como lista de cotejo, lista de frecuencias y escala de estimación.

- Lista de cotejo o de chequeo: también denominada lista de control o de verificación, es un instrumento en el que se indica la presencia o ausencia de un aspecto o conducta a ser observada. Se estructura en tres columnas:

  a) En la columna izquierda se mencionan los elementos o conductas que se pretenden observar.

  b) La columna central dispone de un espacio para marcar en el supuesto de que sea positiva la presencia del aspecto o conducta.

  c) En la columna derecha, se utiliza el espacio para indicar si el elemento o la conducta no está presente.

Ejemplo:

El estudio se propone realizar un diagnóstico sobre la situación actual de la planta física del Colegio Universitario de Caracas (CUC).

Las aulas de clase son amplias	SI ☐	NO ☐
La iluminación es adecuada	SI ☐	NO ☐

La ventilación es suficiente	SI ☐	NO ☐
Los baños funcionan	SI ☐	NO ☐
Existen salidas de emergencia	SI ☐	NO ☐

- Lista de frecuencias: es un instrumento que se diseña para registrar cada vez que se presenta una conducta o comportamiento.

Ejemplo:

CONDUCTA	FRECUENCIA
Intervenciones en clase	☐
Interacción con otros alumnos	☒ ☐
Interrupciones sin justificación	☒ ☐

- Escala de estimación: a diferencia de la lista de cotejo, que sólo considera la presencia o ausencia, este instrumento consiste en una escala que busca medir cómo se manifiesta una situación o conducta.

Ejemplo:

Se realiza una observación al personal de una empresa y se establecen algunos aspectos tales como:

- El trato que se ofrece a los clientes es:

bueno	regular	deficiente
☐	☐	☐

- La presentación personal de los empleados es:

buena	regular	deficiente
☐	☐	☐

- La relación entre los compañeros de trabajo es:

buena	regular	deficiente
☐	☐	☐

6.3. La encuesta y sus instrumentos

Se define la encuesta como una técnica que pretende obtener información que suministra un grupo o muestra de sujetos acerca de sí mismos, o en relación con un tema en particular.

La encuesta puede ser oral o escrita:

La encuesta oral se fundamenta en un interrogatorio "cara a cara" o por vía telefónica, en el cual el encuestador pregunta y el encuestado responde. Contraria a la entrevista, en la encuesta oral se realizan pocas y breves preguntas porque su duración es bastante corta.

Sin embargo, esto permite al encuestador abordar una gran cantidad de personas en poco tiempo. Es decir, la encuesta oral se caracteriza por ser poco profunda, pero de gran alcance. Un ejemplo es el caso de los encuestadores que abordan a las personas en sitios públicos.

Esta modalidad utiliza como instrumento una tarjeta contentiva de las preguntas y opciones de respuesta, la cual siempre es llenada por el encuestador, a diferencia de la encuesta escrita que se realiza a través de un cuestionario autoadministrado, el cual como su nombre lo indica, siempre es respondido de forma escrita por el encuestado.

Otros instrumentos empleados en una encuesta oral, son las grabadoras y cámaras de video, de uso muy común en las encuestas que realizan en la calle los medios de comunicación.

Las encuestas son empleadas frecuentemente en áreas específicas como la investigación de mercado, estudios del consumidor, encuestas electorales y estudios de opinión en general. Los resultados arrojados son utilizados para la toma de decisiones y el diseño de campañas de diversa índole.

Por otra parte, la encuesta escrita es la que se realiza mediante un cuestionario y será tratada más adelante.

6.4. La entrevista y sus instrumentos

La entrevista, más que un simple interrogatorio, es una técnica basada en un diálogo o conversación "cara a cara", entre el entrevistador y el entrevistado acerca de un tema previamente determinado, de tal manera que el entrevistador pueda obtener la información requerida.

Esta técnica se diferencia de la modalidad oral de la encuesta en los siguientes aspectos:

Una entrevista se caracteriza por su profundidad, es decir, indaga de forma amplia en gran cantidad de aspectos y detalles, mientras que la encuesta oral, como se dijo anteriormente, aborda de forma muy precisa o superficial uno o muy pocos aspectos.

Por otra parte, la entrevista tiene un menor alcance en cuanto a la cantidad de personas que pueden ser entrevistadas en un período determinado, es decir, se abarcan menos personas.

Realizar una entrevista puede ocupar un tiempo significativo en un solo entrevistado, a diferencia de la encuesta oral que por su brevedad puede incluir a varias personas en poco tiempo.

La entrevista se clasifica en:

Entrevista estructurada o formal

Es la que se realiza a partir de una guía prediseñada que contiene las preguntas que serán formuladas al entrevistado. En este caso, la misma guía de entrevista puede servir como instrumento para registrar las respuestas, aunque también puede emplearse el grabador o la cámara de video.

Entrevista no estructurada o informal

En esta modalidad no se dispone de una guía de preguntas elaboradas previamente. Sin embargo, se orienta por unos objetivos preestablecidos que permiten definir el tema de la entrevista, de allí que el entrevistador deba poseer una gran habilidad para formular las interrogantes sin perder la coherencia.

Entrevista semi-estructurada

Aun cuando existe una guía de preguntas, el entrevistador puede realizar otras no contempladas inicialmente. Esto se debe a que una respuesta puede dar origen a una pregunta adicional o extraordinaria. Esta técnica se caracteriza por su flexibilidad.

Además de sus instrumentos específicos, tanto la entrevista estructurada como la no estructurada pueden emplear instrumentos tales como el grabador y la cámara de video.

6.5. El cuestionario

Es la modalidad de encuesta que se realiza de forma escrita mediante un instrumento o formato en papel contentivo de una serie de preguntas. Se le denomina cuestionario autoadministrado porque debe ser llenado por el encuestado, sin intervención del encuestador.

Cabe destacar que actualmente, el cuestionario también puede presentarse a través de medios magnéticos (CD o DVD) y electrónicos (correo electrónico e Internet).

El cuestionario puede ser:

a) **Cuestionario de preguntas cerradas:** son aquellas que establecen previamente las opciones de respuesta que puede elegir el encuestado. Éstas se clasifican en: dicotómicas: cuando se ofrecen sólo dos opciones de respuesta; y de selección simple, cuando se ofrecen varias opciones, pero se escoge sólo una.

Ejemplos:

¿Posee usted vivienda propia? SI ☐ NO ☐

En relación con el aborto, ¿usted está?

A favor ☐ En contra ☐

¿Posee usted vehículo propio? SI ☐ NO ☐

En relación con la condena a cadena perpétua, ¿usted está?

A favor ☐ En contra ☐

Ejemplos de preguntas cerradas de selección simple:

¿Cuál es su estado civil?

Soltero ☐
Casado ☐
Viudo ☐

¿Qué música le gusta escuchar?

Pop ☐
Salsa ☐
Rock ☐
Merengue ☐
Romántica ☐
Otra ☐

b) Cuestionario de preguntas abiertas: son las que no ofrecen opciones de respuestas, sino que se da la libertad de responder al encuestado, quien desarrolla su respuesta de manera independiente.

Ejemplo:

¿Qué actividades realizó durante el último cargo desempeñado?

c) Cuestionario mixto: es aquel cuestionario que combina preguntas abiertas, cerradas y mixtas.

6.5.1. Recomendaciones para la elaboración del cuestionario

1. Las preguntas de un cuestionario no se inventan a capricho. Éstas deben tener una correspondencia con los objetivos específicos de la investigación. Además, son producto de la operacionalización de las variables y de la definición de los indicadores. Esto se puede apreciar en los cuadros que aparecen a continuación.

VARIABLE Calidad del servicio en el restaurante		
INDICADORES	PREGUNTAS	OPCIONES DE RESPUESTA
• Trato cortés y amable • Rapidez del servicio • Calidad de la comida	¿El trato recibido en el restaurante fue? ¿El tiempo de servicio fue? ¿Cómo evalúa la comida del restaurante?	• Cortés - indiferente - desagradable • Rápido - normal - lento • Buena - regular - deficiente

VARIABLE Cumplimiento de las funciones universitarias		
DIMENSIONES	INDICADORES	PREGUNTAS
Docencia	N° de secciones asignadas N° de alumnos atendidos	¿Cuántas secciones tiene asignadas? ¿Cuántos alumnos atiende usted de manera formal y continua?
Investigación	N° de investigaciones N° de proyectos	¿Cuántas investigaciones ha realizado durante su trayectoria en la institución? ¿En cuántos proyectos de investigación institucional ha participado?
Extensión	Cursos dictados Actividades extracátedra	¿Ha dictado cursos de extensión universitaria? ¿Ha participado en la conducción de alguna actividad extracátedra relacionada con la institución?

2. Ordenar las preguntas de lo general a lo particular (Ley del embudo). Consiste en colocar al inicio las preguntas más generales y luego las específicas. Ejemplo:

¿Acostumbra usted ver televisión?

SI ☐ NO ☐

Si respondió afirmativamente, indique:

¿En qué horario?

¿Cuál es su canal preferido?

¿Cuál es su programa favorito?

¿Cuál es su artista predilecto?

Otro ejemplo sería:

¿Posee vehículo propio?

SI ☐ NO ☐

De ser positiva su respuesta, señale:
Marca:
Año:
Modelo:
Tipo:
Color:

3. Evitar preguntas que abusen de la memoria del encuestado. Ejemplos:

¿Qué lugar visitó usted durante las vacaciones escolares de 1993?

¿Cuándo fue la última vez que vio una película de cine en blanco y negro?

4. Obviar preguntas sobre temas o conocimientos especializados. Ejemplos:

¿Qué opina usted sobre el desarrollo actual que ha experimentado la nanotecnología?

¿Qué expectativas tiene usted ante el avance de la mecatrónica?

5. No incluir preguntas que induzcan a la respuesta (preguntas guía o sesgadas).

Ejemplo:

Pregunta sesgada:

¿Considera usted que ha sido un fracaso la política económica del gobierno?

Pregunta reformulada:

¿Qué opinión le merece la política económica del gobierno?

6. Omitir las preguntas que originen múltiples interpretaciones.

Ejemplos:

¿Lo ha hecho usted alguna vez?

¿Cuántas veces le gusta hacerlo?

7. Separar las preguntas dobles, es decir, aquellas en las que se interroga sobre dos asuntos en una misma pregunta.

Ejemplo:

¿Le gusta cantar y bailar?

Preguntas reformuladas:

¿Le gusta cantar? SI ☐ NO ☐

¿Le gusta bailar? SI ☐ NO ☐

8. Incluir preguntas que permitan verificar respuestas anteriores o preguntas de control. Es decir, a partir de cualquier pregunta, formule otra interrogante que permita establecer la veracidad de la respuesta emitida anteriormente.

Ejemplo:

¿Conoce usted los principios básicos de la administración?

Si su respuesta es positiva, mencione algunos (Control).

Otra forma de control consiste en repetir una pregunta, pero con una redacción diferente. De esta manera se determina si hay consistencia en las respuestas.

9. Emplear frases de enlace cuando sea necesario.

Ejemplos:

- Si responde afirmativamente, pase a la siguiente...
- De ser negativa su respuesta, diga el porqué...

10. Utilizar escalas de rangos para preguntas sobre asuntos muy personales, tales como, la edad y el salario.

Ejemplo:

Su sueldo mensual se ubica entre:

200.000------ 300.000
300.001------ 400.000
400.001------ 500.000
500.001------ 600.000
más de 600.000

11. Una vez construido el cuestionario, se recomienda aplicar una prueba piloto o sondeo preliminar a un pequeño grupo que no forme parte de la muestra, pero que sea equivalente en cuanto a sus características. Esto con la finalidad de establecer la validez, corregir cualquier falla y elaborar la versión definitiva del instrumento.

La validez del cuestionario significa que las preguntas o ítems deben tener una correspondencia directa con los objetivos de la investigación. Es decir, las interrogantes consultarán sólo aquello que se pretende conocer o medir.

capítulo

9 Recolección de datos cuantitativos

El momento de aplicar los instrumentos de medición y recolectar los datos representa la oportunidad para el investigador de confrontar el trabajo conceptual y de planeación con los hechos.

Roberto Hernández-Sampieri

Proceso de investigación cuantitativa

Paso 8 Recolectar los datos

- Definir la forma idónea de recolectar los datos de acuerdo con el planteamiento del problema y las etapas previas de la investigación.
- Seleccionar o elaborar uno o varios instrumentos o métodos para recolectar los datos requeridos.
- Aplicar los instrumentos o métodos.
- Obtener los datos.
- Codificar los datos.
- Archivar los datos y prepararlos para su análisis por computadora.

Objetivos de aprendizaje

Al terminar este capítulo, el alumno será capaz de:

1. Visualizar diferentes métodos para recolectar datos cuantitativos.
2. Entender el significado de "medir" y su importancia en el proceso cuantitativo.
3. Comprender los requisitos que toda recolección de datos debe incluir.
4. Conocer los principales instrumentos para recolectar datos cuantitativos.
5. Elaborar y aplicar los diferentes instrumentos de recolección de datos cuantitativos.
6. Preparar los datos para su análisis cuantitativo.

Síntesis

En el capítulo se analizan los requisitos que un instrumento de medición debe cubrir para recolectar apropiadamente datos cuantitativos: confiabilidad, validez y objetividad. Asimismo, se define el concepto de medición y los errores que pueden cometerse al recolectar datos.

También se explica el proceso para elaborar un instrumento de medición y las principales alternativas para recolectar datos: cuestionarios y escalas de actitudes. Por último, se examina el procedimiento de codificación de datos cuantitativos y la forma de prepararlos para su análisis. Otros temas relacionados con la recolección de los datos se incluyen en el capítulo 5 del centro de recursos en línea.

Centro de recursos en línea

Fases de construcción de un instrumento:

1. Redefiniciones fundamentales
2. Revisión enfocada de la literatura en instrumentos pertinentes
3. Identificación del dominio de las variables a medir y sus indicadores
4. Toma de decisiones clave
5. Construcción del instrumento
6. Prueba piloto
7. Elaboración de la versión final del instrumento o sistema y su procedimiento de aplicación
8. Entrenamiento del personal que administrará el instrumento, codificará respuestas e interpretará resultados
9. Obtener autorizaciones para aplicar el instrumento
10. Administración del instrumento

Recolección de datos cuantitativos

se realiza mediante

Instrumento(s) de medición

Debe(n) representar verdaderamente la(s) variable(s) de la investigación

sus requisitos son:

Confiabilidad

Grado en que un instrumento produce resultados consistentes y coherentes

Procedimientos para determinar la confiabilidad:

- Medida de estabilidad
- Método de formas alternativas o paralelas
- Método de mitades partidas
- Medidas de consistencia interna

Validez

Grado en que un instrumento mide la variable que pretende medir

De ella derivan distintos tipos de evidencia:

- Validez de contenido
- Validez de criterio
- Validez de constructo

Validez total es la consideración de los tipos de evidencia

Objetividad

Grado en que el instrumento es permeable a los sesgos y tendencias del investigador que lo administra, califica e interpreta

Cuyas respuestas se obtienen, codifican y transfieren a una matriz o base de datos y se preparan para su análisis mediante un paquete estadístico para computadora

Tipos

Cuestionarios

- Se basan en preguntas que pueden ser cerradas o abiertas
- Sus contextos pueden ser: autoadministrados, entrevista personal o telefónica, vía internet

Escalas de medición de actitudes que pueden ser:

- Escalamiento tipo Likert
- Diferencial semántico
- Escalograma de Guttman (en el centro de recursos en línea)

Otros tipos son (los encontrará en el centro de recursos en línea)

- Analisis de contenido cuantitativo
- Observación
- Pruebas estandarizadas e inventarios
- Datos secundarios (recolectados por otros investigadores)
- Aparatos y equipos
- Indicadores

Nota: El capítulo se termina de integrar con otro que puede descargarse del centro de recursos en línea en: Material complementario → Capítulos → Capítulo 7, "Recolección de los datos cuantitativos: segunda parte", que contiene alternativas de instrumentos para recolectar datos como el análisis de contenido y los sistemas de observación, además de pruebas e inventarios, escalograma de Guttman (escala de actitudes), datos secundarios e indicadores.

Centro de recursos en línea

¿Qué implica la etapa de recolección de datos?

1 Una vez que seleccionamos el diseño de investigación apropiado y la muestra adecuada de acuerdo con nuestro problema de estudio e hipótesis (si es que se establecieron), la siguiente etapa consiste en recolectar los datos pertinentes sobre los atributos, conceptos o variables de las unidades de muestreo/análisis o casos (participantes, grupos, fenómenos, procesos, organizaciones, etcétera).

Recolectar los datos implica elaborar un plan detallado de procedimientos que nos conduzcan a reunir datos con un propósito específico. Este plan incluye determinar:

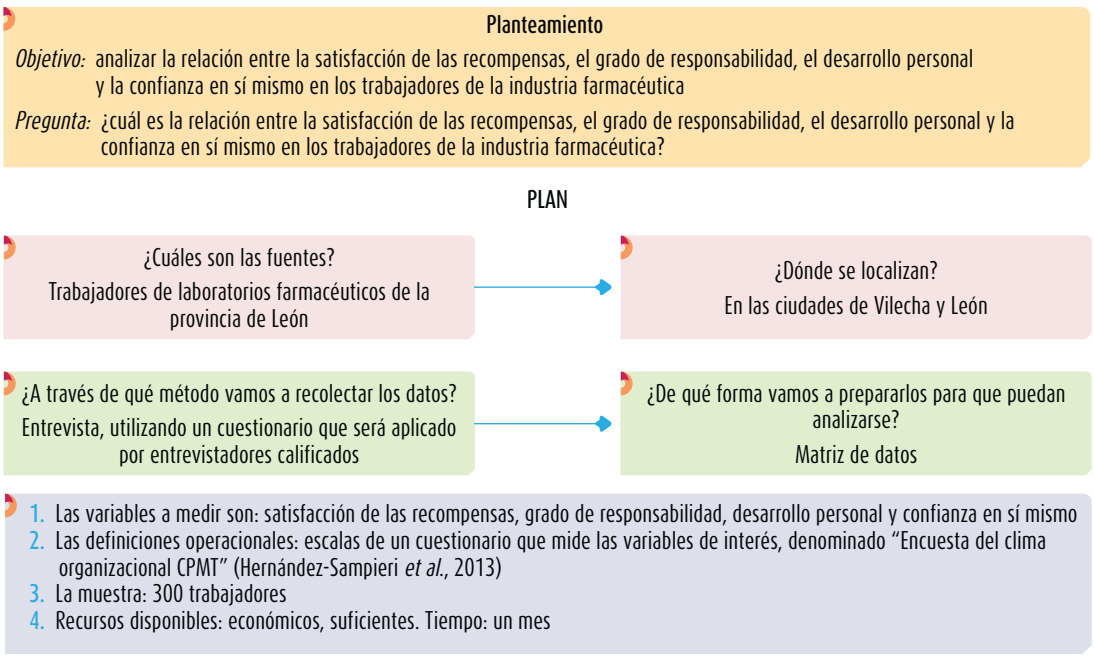
a) ¿Cuáles son las fuentes de las que se obtendrán los datos? Es decir, los datos van a ser proporcionados por personas, se producirán de observaciones y registros o se encuentran en documentos, archivos, bases de datos, etcétera.
b) ¿En dónde se localizan tales fuentes? Regularmente en la muestra seleccionada, pero es indispensable definir con precisión.
c) ¿A través de qué medio o método vamos a recolectar los datos? Esta fase implica elegir uno o varios medios y definir los procedimientos que utilizaremos en la recolección de los datos. El método o métodos deben ser confiables, válidos y "objetivos".
d) Una vez recolectados, ¿de qué forma vamos a prepararlos para que puedan analizarse y respondamos al planteamiento del problema?

El plan se nutre de diversos elementos:

1. Las *variables*, *conceptos* o *atributos* a medir (contenidos en el planteamiento e hipótesis o directrices del estudio).
2. Las *definiciones operacionales*. La manera como hemos operacionalizado las variables es crucial para determinar el método para medirlas, lo cual a su vez, resulta fundamental para realizar las inferencias de los datos.
3. La *muestra*.
4. Los *recursos disponibles* (de tiempo, apoyo institucional, económicos, etcétera).

El plan se implementa para obtener los datos requeridos, no olvidemos que todos los atributos, cualidades y variables deben ser medibles. En la figura 9.1 se encuentra un ejemplo de un plan de este tipo.

Figura 9.1 Ejemplo de plan para la obtención de datos.



www.elosopanda.com | jamespoetrodriguez.com

Con la finalidad de recolectar datos disponemos de una gran variedad de instrumentos o técnicas, tanto cuantitativas como cualitativas, es por ello que en un mismo estudio podemos utilizar ambos tipos. Incluso, hay instrumentos como la Prueba de Propósito Vital (PIL) (que evalúa el propósito de vida de una persona) de Crumbaugh y Maholick (1969) que contiene una parte cuantitativa y una cualitativa (Thoits, 2012; Marsh, Smith, Piek y Saunders, 2003; y Brown, Ashcroft y Miller, 1998). Esto se revisará en el capítulo 17: "Los métodos mixtos".

Por ahora, nos concentraremos en la recolección de los datos cuantitativos y sus conceptos esenciales.

¿Qué significa medir?

2

En la vida diaria medimos constantemente. Por ejemplo, al levantarnos por las mañanas, miramos el reloj despertador y "medimos" la hora; al bañarnos, *ajustamos* la temperatura del agua en la tina o la regadera, *calculamos* la cantidad de café que habremos de servir en la cafetera; nos asomamos por la ventana y estimamos cómo será el día para decidir la ropa que nos pondremos. Al ver el tráfico evaluamos e inferimos a qué hora llegaremos a la universidad o al trabajo, así como la velocidad a la que transitamos (o bien observamos el velocímetro); en ocasiones contamos cuántos anuncios espectaculares observamos en el trayecto u otras cuestiones, incluso deducimos, a partir de ciertos signos, acerca del operador del autobús y de otros conductores: ¿qué tan alegres o enojados están? Como esas, hacemos muchas otras actividades. Medir es parte de nuestras vidas (Bostwick y Kyte, 2005).

En concordancia con la definición clásica del término, ampliamente difundida, **medir** significa "asignar números, símbolos o valores a las propiedades de objetos o eventos de acuerdo con reglas" (Stevens, 1951). Desde luego, no se asignan a los objetos, sino a sus propiedades (Bostwick y Kyte, 2005). Sin embargo, como señalan Carmines y Zeller (1991), esta definición es más apropiada para las ciencias físicas que para las ciencias sociales, ya que varios de los fenómenos que son medidos en éstas no pueden caracterizarse como objetos o eventos, son demasiado abstractos para ello. La disonancia cognitiva, la pareja ideal, el clima organizacional, la cultura fiscal y la credibilidad son conceptos tan abstractos que no deben ser considerados "cosas que pueden verse o tocarse" (definición de objeto) ni solamente como "resultado, consecuencia o producto" (definición de evento) (Carmines y Zeller, 1991). Este razonamiento nos lleva a proponer que es más adecuado definir la **medición** como "el proceso de vincular conceptos abstractos con indicadores empíricos", el cual se realiza mediante un plan explícito y organizado para clasificar (y con frecuencia cuantificar) los datos disponibles (los indicadores), en términos del concepto que el investigador tiene en mente (Carmines y Zeller, 1991). En este proceso, el instrumento de medición o de recolección de datos tiene un papel central. Sin él, no hay observaciones clasificadas.

Medición Proceso que vincula conceptos abstractos con indicadores empíricos.

La definición sugerida incluye dos consideraciones: la primera es desde el punto de vista empírico y se resume en que el centro de atención es la respuesta observable (sea una alternativa de respuesta marcada en un cuestionario, una conducta registrada mediante observación, un valor de un instrumento y su interpretación o una respuesta dada a un entrevistador). La segunda es desde una perspectiva teórica y se refiere a que el interés se sitúa en el concepto subyacente no observable que se representa por medio de la respuesta. Así, los registros del instrumento de medición representan valores visibles de conceptos abstractos. Un **instrumento de medición** adecuado es aquel que registra datos observables que representan verdaderamente los conceptos o las variables que el investigador tiene en mente (Grinnell, Williams y Unrau, 2009). En términos cuantitativos: capturo verdaderamente la "realidad" que deseo capturar. Bostwick y Kyte (2005) lo señalan de la siguiente forma: la función de la medición es establecer una correspondencia entre el "mundo real" y el "mundo conceptual". El primero provee evidencia empírica, el segundo proporciona modelos teóricos para encontrar sentido a ese segmento del mundo real que estamos tratando de describir.

Instrumento de medición Recurso que utiliza el investigador para registrar información o datos sobre las variables que tiene en mente.

En toda investigación cuantitativa aplicamos un instrumento para medir las variables contenidas en las hipótesis (y cuando no hay hipótesis simplemente para medir las variables de interés). Esa

medición es eficaz cuando el instrumento de recolección de datos en realidad representa las variables que tenemos en mente. Si no es así, nuestra medición es deficiente; por tanto, la investigación no es digna de tomarse en cuenta. Desde luego, no hay medición perfecta. Es casi imposible que representemos con fidelidad variables como la inteligencia emocional, la motivación, el nivel socioeconómico, el liderazgo democrático, el abuso sexual infantil y otras más; pero es un hecho que debemos acercarnos lo más posible a la representación fiel de las variables que se van a observar mediante el instrumento de medición que desarrollemos. Se trata de un precepto básico del enfoque cuantitativo. Al medir estandarizamos y cuantificamos los datos (Babbie, 2014 y Bostwick y Kyte, 2005).

¿Qué requisitos debe cubrir un instrumento de medición?

3 Toda medición o instrumento de recolección de datos debe reunir tres requisitos esenciales: *confiabilidad, validez* y *objetividad.*

Confiabilidad

Confiabilidad Grado en que un instrumento produce resultados consistentes y coherentes.

La **confiabilidad** de un instrumento de medición se refiere al grado en que su aplicación repetida al mismo individuo u objeto produce resultados iguales (Hernández-Sampieri *et al.,* 2013; Kellstedt y Whitten, 2013; y Ward y Street, 2009). Por ejemplo, si se midiera en este momento la temperatura ambiental usando un termómetro y éste indicara que hay 22°C, y un minuto más tarde se consultara otra vez y señalara 5°C, tres minutos después se observara nuevamente y éste indicara 40°C, dicho termómetro no sería confiable, ya que su aplicación repetida produce resultados distintos. Asimismo, si una prueba de inteligencia (*Intelligence Quotient*, IQ) se aplica hoy a un grupo de personas y da ciertos valores de inteligencia, se aplica un mes después y proporciona valores diferentes, al igual que en subsecuentes mediciones, tal prueba no sería confiable (analice los valores de la tabla 9.1, suponiendo que los coeficientes de inteligencia oscilaran entre 100 y 135). Los resultados no son coherentes, pues no se puede "confiar" en ellos.

Tabla 9.1 Ejemplo de resultados proporcionados por un instrumento de medición sin confiabilidad

Primera aplicación		Segunda aplicación		Tercera aplicación	
Mariana	135	Sergio	131	Guadalupe	127
Viridiana	125	Laura	130	Guadalupe	127
Sergio	118	Chester	125	Mariana	118
Laura	110	Guadalupe	112	Laura	115
Guadalupe	108	Mariana	110	Chester	112
Chester	106	Viridiana	105	Viridiana	108
Agustín	100	Agustín	101	Sergio	105

La confiabilidad de un instrumento de medición se determina mediante diversas técnicas, las cuales se comentarán brevemente después de revisar los conceptos de validez y objetividad.

Validez

Validez Grado en que un instrumento en verdad mide la variable que se busca medir.

La **validez**, en términos generales, se refiere al grado en que un instrumento mide realmente la variable que pretende medir. Por ejemplo, un instrumento válido para medir la inteligencia debe medir la inteligencia y no la memoria. Un método para medir el rendimiento bursátil tiene que medir precisamente esto y no la imagen de una empresa. Un ejemplo —aunque muy obvio— de completa invalidez sería intentar medir el peso de los objetos con una cinta métrica en lugar de con una báscula.

En apariencia es sencillo lograr la validez. Después de todo, como dijo un estudiante: "pensamos en la variable y vemos cómo hacer preguntas o imaginar indicadores sobre esa variable". Esto resultaría factible en unos cuantos casos (como lo sería el género al que pertenece una persona). Sin embargo, la situación no es tan simple cuando se trata de variables como la motivación, la calidad del servicio a los clientes, la actitud hacia un candidato político y menos aún con sentimientos y emociones, así como de otras variables con las que trabajamos en todas las ciencias. La validez es una cuestión más compleja que debe alcanzarse en todo instrumento de medición que se aplica. Kerlinger (1979, p. 138) plantea la siguiente pregunta respecto de la validez: ¿está midiendo lo que cree que está midiendo? Si es así, su medida es válida; si no, evidentemente carece de validez.

La validez es un concepto del cual pueden tenerse diferentes tipos de evidencia (Babbie, 2014; Hays, 2013; Kellstedt y Whitten, 2013; The SAGE Glossary of the Social and Behavioral Sciences, 2009a; Streiner y Norman, 2008; Wiersma y Jurs, 2008; Gallestey, 2007; Rupp y Pant, 2006; Carmines y Woods, 2003a; y Gronlund, 1990): 1) *evidencia relacionada con el contenido,* 2) *evidencia relacionada con el criterio* y 3) *evidencia relacionada con el constructo.* A continuación analizaremos cada una de ellas.

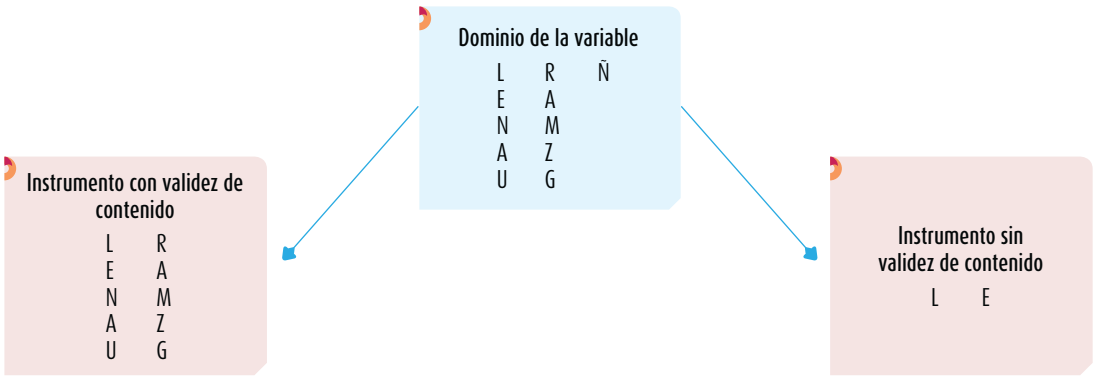
1. Evidencia relacionada con el contenido

La **validez de contenido** se refiere al grado en que un instrumento refleja un dominio específico de contenido de lo que se mide. Es el grado en el que la medición representa al concepto o variable medida (The SAGE Glossary of the Social and Behavioral Sciences, 2009b y Bohrnstedt, 1976). Por ejemplo, una prueba de operaciones aritméticas no tendrá validez de contenido si incluyera sólo problemas de resta y excluyera problemas de suma, multiplicación o división. O bien, una prueba de conocimientos sobre las canciones de Los Beatles no deberá basarse solamente en sus álbumes *Let it Be* y *Abbey Road*, sino que debe incluir canciones de todos sus discos. O una prueba de conocimientos de líderes históricos de América Latina que omita a Simón Bolívar, Salvador Allende, Túpac Amaru II, Ernesto *el Che* Guevara, José de San Martín o Benito Juárez.

Validez de contenido Grado en que un instrumento refleja un dominio específico de contenido de lo que se mide.

Un instrumento de medición requiere tener representados a todos o la mayoría de los componentes del dominio de contenido de las variables que se van a medir. Este hecho se ilustra en la figura 9.2.

Figura 9.2 Ejemplo de un instrumento de medición con validez de contenido frente a otro que carece de ella.



El dominio de contenido de una variable normalmente está definido o establecido por la literatura (teoría y trabajos antecedentes). En indagaciones exploratorias en las que las fuentes previas son escasas, el investigador comienza a adentrarse en el problema de estudio y a proponer cómo puede estar constituido tal dominio. De cualquier manera, en cada investigación uno debe probar que el instrumento utilizado es válido. Un ejemplo del intento por establecer el dominio de contenido de una variable es el siguiente:

Ejemplo

Hernández-Sampieri (2005), para establecer el dominio de la variable clima organizacional, revisó 20 estudios clásicos sobre el concepto, comprendidos entre 1964 y 1977, así como más de 100 investigaciones publicadas en revistas científicas entre 1975 y 2005. Por otro lado, consideró diversos libros sobre el tema, tres metaanálisis y otras tantas revisiones del estado del conocimiento sobre dicho clima. También evaluó 15 estudios efectuados en el contexto en el que habría de llevar a cabo su propia investigación. Encontró que en la literatura se han considerado decenas de dimensiones o componentes del clima organizacional, por lo que realizó un análisis para determinar cuáles habían sido los más frecuentes, éstos fueron: 1) moral, 2) apoyo de la dirección, 3) innovación, 4) identificación con la empresa, 5) comunicación, 6) percepción del desempeño, 7) motivación intrínseca, 8) autonomía, 9) satisfacción general, 10) liderazgo, 11) visión y 12) recompensas o retribución. Dejó a un lado otros, como confianza en sí mismo y estándares de excelencia o conformidad. Con lo anterior generó su instrumento de medición.

Si el dominio de un instrumento es demasiado estrecho con respecto al dominio de la variable, el primero no la representará. La pregunta que se responde con la **validez de contenido** es: ¿el instrumento mide adecuadamente las principales dimensiones de la variable en cuestión? En un cuestionario, por ejemplo, cabría interrogar: ¿qué tan bien representan las preguntas a todas las que pudieran hacerse?

2. Evidencia relacionada con el criterio

Validez de criterio Validez que se establece al correlacionar las puntuaciones resultantes de aplicar el instrumento con las puntuaciones obtenidas de otro criterio externo que pretende medir lo mismo.

La **validez de criterio** de un instrumento de medición se establece al comparar sus resultados con los de algún criterio externo que pretende medir lo mismo. Supongamos que Fernando trata de "medir" el grado en que es aceptado por Laura. Entonces decide que va a tomarla de la mano y observará su reacción. Supuestamente, si ella no retira la mano, esto indicaría cierta aceptación. Pero para asegurarse de que su medición es válida, decide utilizar otra forma de medición adicional, por ejemplo, mirarla fijamente sin apartar la vista de sus ojos. En apariencia, si Laura le sostiene la mirada, esto sería otro indicador de aceptación. Así, su medición de aceptación se valida mediante dos métodos al comparar dos criterios. El ejemplo tal vez sea simple, pero describe la esencia de la validez relativa al criterio.

Este criterio es un estándar con el que se juzga la validez del instrumento (Jackson, 2011 y The SAGE Glossary of the Social and Behavioral Sciences, 2009c). Cuanto más se relacionen los resultados del instrumento de medición con los del criterio, la validez será mayor. Por ejemplo, un investigador valida un examen sobre manejo de aviones al mostrar la exactitud con la que el examen predice qué tan bien un grupo de pilotos es capaz de operar un aeroplano.

Si el criterio se fija en el presente de manera paralela, se habla de **validez concurrente** (Kaplan y Saccuzzo, 2013) (los resultados del instrumento se correlacionan con el criterio en el mismo momento o punto de tiempo). Por ejemplo, Núñez (2001) desarrolló una herramienta para medir el sentido de vida de acuerdo con las ideas de Viktor Frankl, el test Celaya. Para aportar evidencia de validez en relación con su instrumento, lo aplicó y a su vez administró otros instrumentos que miden conceptos muy similares, tal como la PIL (Prueba de Propósito Vital) de Crumbaugh y Maholick (1969) y el Logo Test de Lukas (1984). Luego comparó las puntuaciones de los participantes en las tres pruebas, demostró que las correlaciones entre las puntuaciones eran significativamente elevadas. De esta manera fue como aportó validez concurrente para su instrumento.

Si el criterio se fija en el futuro, se habla de **validez predictiva** (The SAGE Glossary of the Social and Behavioral Sciences, 2009d). Por ejemplo, una prueba para determinar la capacidad gerencial de candidatos a ocupar altos puestos ejecutivos se validaría comparando sus resultados con el desempeño posterior de los ejecutivos en su trabajo regular. Un cuestionario para detectar las preferencias del electorado por los distintos partidos contendientes y por sus candidatos en la época de las campañas, puede validarse comparando sus resultados con los resultados finales y definitivos de la elección.

El principio de la validez de criterio es sencillo: si diferentes instrumentos o criterios miden el mismo concepto o variable, deben arrojar resultados similares. Bostwick y Kyte (2005) lo expresan de la siguiente forma:

Si hay validez de criterio, las puntuaciones obtenidas por ciertos casos en un instrumento deben estar correlacionadas y predecir las puntuaciones de estos mismos casos logradas en otro criterio.

La pregunta que se responde con la validez de criterio es: ¿en qué grado el instrumento comparado con otros criterios externos mide lo mismo?

3. Evidencia relacionada con el constructo

La **validez de constructo** es probablemente la más importante, sobre todo desde una perspectiva científica, y se refiere a qué tan bien un instrumento representa y mide un concepto teórico (Babbie, 2014; Grinnell, Williams y Unrau, 2009; The SAGE Glossary of the Social and Behavioral Sciences, 2009d; y Sawilowsky, 2006). A esta validez le concierne en particular el significado del instrumento, esto es, qué está midiendo y cómo opera para medirlo. Integra la evidencia que soporta la interpretación del sentido que poseen las puntuaciones del instrumento (Messick, 1995).

Validez de constructo Debe explicar cómo las mediciones del concepto o variable se vinculan de manera congruente con las mediciones de otros conceptos correlacionados teóricamente.

Parte del grado en el que las mediciones del concepto proporcionadas por el instrumento se relacionan de manera consistente con mediciones de otros conceptos o variables vinculadas empírica y teóricamente (por la teoría, modelos e hipótesis previas o investigaciones antecedentes). A tales conceptos se les denomina "constructos". Un **constructo** es una variable medida y tiene lugar dentro de una hipótesis, teoría o modelo teórico. Es un atributo que no existe aislado sino en relación con otros y debe ser inferido de la evidencia que tenemos en nuestras manos y que proviene de las puntuaciones del instrumento aplicado.

Constructo o construcción Variable medida que tiene lugar dentro de una hipótesis, teoría o modelo teórico.

La validez de constructo incluye tres etapas (Carmines y Zeller, 1991):

1. Sobre la base de la revisión de la literatura, se establece y especifica la relación entre el concepto o variable medida por el instrumento y los demás conceptos incluidos en la teoría, modelo teórico o hipótesis.
2. Se asocian estadísticamente los conceptos y se analizan cuidadosamente las correlaciones.
3. Se interpreta la evidencia empírica de acuerdo con el nivel en el que se clarifica la validez de constructo de una medición en particular.

El proceso de *validación de un constructo* está vinculado con la teoría. No es conveniente llevar a cabo tal validación, a menos que exista un marco teórico que soporte la variable en relación con otras variables. Desde luego, no es necesaria una teoría muy desarrollada, pero sí investigaciones que hayan demostrado que los conceptos se asocian. Cuanto más elaborada y comprobada se encuentre la teoría que apoya la hipótesis, la validación del constructo arrojará mayor luz sobre la validez general de un instrumento de medición. Tenemos más confianza en la validez de constructo de una medición cuando sus resultados se correlacionan significativamente con un mayor número de mediciones de variables que, en teoría y de acuerdo con estudios antecedentes, están relacionadas. Veamos la validez de constructo con el ejemplo ya comentado sobre el clima organizacional.

Ejemplo

En el caso de Hernández-Sampieri (2005), que como recordamos, desarrolló un instrumento para evaluar al clima organizacional considerando 12 variables: moral, apoyo de la dirección, innovación, etc., la pregunta obvia es: ¿tal instrumento realmente mide el clima organizacional? ¿Verdaderamente lo representa? En cuanto a contenido se demostró que sí reflejaba las principales dimensiones del clima organizacional. Pero esto no es suficiente, necesita demostrar que su instrumento concuerda con la teoría. Ésta, basada en diversos estudios, indica que tales dimensiones se encuentran fuertemente vinculadas y que se unen o "funden" entre sí para formar un constructo multidimensional denominado clima organizacional, y que además se asocian con el involucramiento en el trabajo y el compromiso organizacional. Entonces, para aportar validez de constructo, se correlacionaron todas las dimensiones entre sí y luego la escala de clima con dicho involucramiento y compromiso. Tales vínculos se encontraron mediante análisis estadístico y los resultados coincidieron con la teoría y se obtuvo evidencia sobre la validez de constructo del instrumento.

Las preguntas que se responden con la validez de constructo son: ¿el concepto teórico está realmente reflejado en el instrumento? ¿Qué significan las puntuaciones del instrumento? ¿El instrumento mide el constructo y sus dimensiones? ¿Por qué? ¿Cómo opera el instrumento?

Validez de expertos Grado en que un instrumento realmente mide la variable de interés, de acuerdo con expertos en el tema.

Otro tipo de validez que algunos autores consideran es la **validez de expertos** o *face validity*, la cual se refiere al grado en que aparentemente un instrumento mide la variable en cuestión, de acuerdo con "voces calificadas". Se encuentra vinculada a la validez de contenido y, de hecho, se consideró por muchos años como parte de ésta. Hoy se concibe como un tipo adicional de evidencia (Gravetter y Forzano, 2011; Streiner y Norman, 2008; y Mostert, 2006). Regularmente se establece mediante la evaluación del instrumento ante expertos. Por ejemplo, Hernández-Sampieri (2005) sometió el instrumento a revisión por parte de asesores en desarrollo organizacional, académicos y gerentes de recursos humanos. Asimismo, más recientemente se ha hablado de la *validez consecuente*, que se refiere a las secuelas sociales del uso e interpretación de una prueba (Mertens, 2010).

Validez total

La validez de un instrumento de medición se evalúa sobre la base de todos los tipos de evidencia. Cuanta mayor evidencia de validez de contenido, de validez de criterio y de validez de constructo tenga un instrumento de medición, éste se acercará más a representar las variables que pretende medir.

Validez total = validez de contenido + validez de criterio + validez de constructo

Relación entre la confiabilidad y la validez

Un instrumento de medición puede ser confiable, pero no necesariamente válido (un aparato, por ejemplo, quizá sea consistente en los resultados que produce, pero puede no medir lo que pretende).

La validez y la confiabilidad no se asumen, se prueban.

Por ello es requisito que el instrumento de medición demuestre ser **confiable y válido**. De no ser así, los resultados de la investigación no deben tomarse en serio.

Para ampliar este comentario, recurriremos a la analogía de Bostwick y Kyte (2005, pp. 108-109). Supongamos que vamos a probar un arma con tres tiradores. Cada uno debe realizar cinco disparos, entonces:

- Tirador 1 Sus disparos no impactan en el centro del blanco y se encuentran diseminados por todo el blanco.
- Tirador 2 Tampoco impacta en el centro del blanco, aunque sus disparos se encuentran cercanos entre sí, fue consistente, mantuvo un patrón.
- Tirador 3 Los disparos se encuentran cercanos entre sí e impactaron en el centro del blanco.

Sus resultados podrían visualizarse como en la figura 9.3, en la cual se vinculan la confiabilidad y la validez.

Figura 9.3 Representación de la confiabilidad y la validez.

Factores que pueden afectar la confiabilidad y la validez

Hay diversos factores que llegan a afectar la confiabilidad y la validez de los instrumentos de recolección de los datos e introducen errores en la medición.[1] A continuación se mencionarán los más comunes.

La improvisación. Algunas personas creen que elegir un instrumento de medición o desarrollarlo es algo que puede tomarse a la ligera. Incluso, ciertos profesores piden a los alumnos que construyan instrumentos de medición de un día para otro o, lo que es igual, de una semana a otra, lo cual habla del poco o nulo conocimiento del proceso de elaboración de instrumentos de recolección de los datos. Esta improvisación genera casi siempre instrumentos poco válidos o confiables.

Aun a los investigadores experimentados les toma tiempo desarrollar un instrumento de medición. Además, para construirlo se requiere conocer muy bien la variable que se pretende medir, así como la teoría y la práctica que la sustentan.

Cuando las mediciones se refieren a seres humanos hay otros factores:

a) *Utilizar instrumentos desarrollados en el extranjero que no han sido validados en nuestro contexto*: cultura y tiempo. Traducir un instrumento, aun cuando adaptemos los términos a nuestro lenguaje y los contextualicemos, no es ni remotamente una validación. Constituye un primer y necesario paso, aunque sólo es el principio. En el caso de traducciones, es importante verificar que los términos centrales tengan referentes con el mismo significado —o alguno muy parecido— en la cultura en la que se va a utilizar dicho instrumento (vincular términos entre la cultura de origen y la cultura destinataria). A veces se traduce, se obtiene una versión y ésta, a su vez, se vuelve a traducir al idioma original (traducción inversa).

Por otra parte, existen instrumentos que fueron validados en nuestro contexto, pero hace mucho tiempo. Hay instrumentos en los que hasta el lenguaje "nos suena anticuado". Las culturas, los grupos y las personas cambian; y esto debemos tomarlo en cuenta al elegir o desarrollar un instrumento de medición.

b) *Usar instrumentos inadecuados para las personas a quienes se les aplica*: no son empáticos. Utilizar un lenguaje muy elevado para los sujetos respondientes, no tomar en cuenta diferencias de género, edad, conocimientos, memoria, nivel ocupacional y educativo, motivación para contestar, capacidades de conceptualización y otras diferencias en los participantes, son errores que llegan a afectar la validez y la confiabilidad. Este error ocurre a menudo cuando los instrumentos deben administrarse a niños. Asimismo, hay grupos de la población que requieren instrumentos apropiados para ellos, tal es el caso de las personas con capacidades distintas. En la actualidad se han desarrollado diversas pruebas que las toman en cuenta (por ejemplo, pruebas en sistema Braille para individuos con capacidades visuales distintas o pruebas orales para personas que no pueden escribir). Otro ejemplo son los indígenas o inmigrantes de otras culturas, pues en ocasiones se les aplican instrumentos en un idioma o contexto que no es el suyo.

Quien realiza una investigación debe adaptarse siempre a los participantes y no al revés, ya que es necesario brindarles todo tipo de facilidades. Si éste es el caso, se sugiere consultar a Mertens y McLaughlen (2004), en cuyo libro hay un capítulo dedicado a la recolección de información de personas con capacidades diferentes o de culturas especiales, y a Eckhardt y Anastas (2007). Asimismo, es recomendable revisar la página web de alguna asociación internacional como la American Psychological Association.

c) *Cuestiones vinculadas con los estilos personales de los participantes* (Bostwick y Kyte, 2005) como: deseabilidad social (tratar de dar una impresión muy favorable a través de las respuestas), tendencia a asentir con respecto a todo lo que se pregunta, dar respuestas inusuales o contestar siempre negativamente.

[1] Se ha omitido intencionalmente la exposición de los errores sistemáticos y no sistemáticos que afectan la confiabilidad y la validez, con objeto de simplificar al lector las explicaciones. Un comentario se incluye en el centro de recursos en línea en: Material complementario → Capítulos → Capítulo 7 "Recolección de los datos cuantitativos: segunda parte".

Centro de recursos en línea

d) *Condiciones en las que se aplica el instrumento de medición.* El ruido, la inadecuada iluminación, el frío (por ejemplo, en una encuesta de casa en casa), un instrumento demasiado largo o tedioso, una encuesta telefónica después de que algunas compañías han utilizado el mercadeo telefónico en exceso y a destiempo (promocionar servicios a las 7 a.m. de un domingo o después de las 11 p.m. entre semana) son cuestiones que llegan a afectar negativamente la validez y la confiabilidad, al igual que si el tiempo que se brinda para responder al instrumento es inapropiado. Por lo común en los experimentos se cuenta con instrumentos de medición más largos y complejos que en los diseños no experimentales. Por ejemplo, en una encuesta pública sería muy difícil aplicar una prueba larga o compleja.

Otro factor que se encuentra en todo tipo de instrumentos en cualquier campo de conocimiento es *la falta de estandarización*: que las instrucciones no sean las mismas para todos los participantes, que el orden de las preguntas sea distinto para algunos individuos, que los instrumentos de observación no resulten equivalentes, que el procedimiento para administrar la medición no sea exactamente el mismo para todos los casos (por ejemplo, al medir la presión arterial, que a unos pacientes se les tome después de un tiempo de relajación y a otros, no; o bien, que al pesar ciertas piezas no se utilice la misma báscula). Este elemento también se vincula con la objetividad.

Aspectos mecánicos como que si el instrumento es escrito, que no sean legibles las instrucciones, falten páginas, no haya espacio adecuado para contestar o no se comprendan las instrucciones, también influyen de manera desfavorable. Lo mismo pasaría con un aparato mal calibrado.

Con respecto a la validez de constructo dos factores pueden afectarla significativamente:

a) La estrechez del contenido, es decir, que se excluyan dimensiones importantes de la variable o las variables medidas.

b) La amplitud exagerada, donde el riesgo es que el instrumento contenga excesiva intrusión de otros constructos similares.

Muchos de los errores se pueden evitar mediante una adecuada revisión de la literatura, que nos permite seleccionar las dimensiones apropiadas de las variables del estudio, otros criterios para comparar los resultados de nuestro instrumento, teorías de respaldo, instrumentos de dónde elegir, etcétera.

Objetividad

Es difícil lograr la objetividad, particularmente en el caso de las ciencias sociales. En ciertas ocasiones se alcanza mediante el consenso o mediciones múltiples. Al tratarse de cuestiones físicas las percepciones suelen compartirse (por ejemplo, la mayoría de las personas estarían de acuerdo en que el agua de mar contiene sal o los rayos del Sol queman), pero en temas que tienen que ver con la conducta humana como los valores, las atribuciones y las emociones, el consenso es más complejo. Imaginemos que 10 observadores deben ver una película y calificarla como "muy violenta", "violenta", "neutral", "poco violenta" y "nada violenta". Tres personas indican que es muy violenta, tres que es violenta y cuatro la evalúan como neutral; es difícil responder qué tan violenta es la película. O bien, ¿quién fue mejor compositor: Mozart, Beethoven o Bach? Todo es relativo. Sin embargo, la objetividad aumenta al reducirse la incertidumbre (Unrau, Grinnell y Williams, 2011). Desde luego, la certidumbre total no existe en ninguna ciencia; el conocimiento es aceptado como verdadero, hasta que nueva evidencia demuestra lo contrario.

Objetividad del instrumento Grado en que el instrumento es o no permeable a la influencia de los sesgos y tendencias de los investigadores que lo administran, califican e interpretan.

En un instrumento de medición, la **objetividad** se refiere al grado en que éste es o no permeable a la influencia de los sesgos y tendencias del investigador o investigadores que lo administran, califican e interpretan (Mertens, 2010). Investigadores racistas o machistas quizás influyan negativamente por su sesgo contra un grupo étnico o el género femenino. Lo mismo podría suceder con las tendencias ideológicas, políticas, religiosas o la orientación sexual. En este sentido, los aparatos y sistemas calibrados (por ejemplo, una pistola láser para medir la velocidad de un automóvil) son más objetivos que otros sistemas que requieren cierta interpretación (como un detector de mentiras) y éstos, a su vez, más objetivos que las pruebas estandarizadas, las cuales son menos subjetivas que las pruebas proyectivas.

La objetividad se refuerza mediante la estandarización en la aplicación del instrumento (mismas instrucciones y condiciones para todos los participantes) y en la evaluación de los resultados; así como al emplear personal capacitado y experimentado en el instrumento. Por ejemplo, si se utilizan observadores, su proceder en todos los casos debe ser lo más similar posible y su entrenamiento tendrá que ser profundo y adecuado.

Los estudios cuantitativos buscan que la influencia de las características y las tendencias del investigador se reduzca al mínimo posible, lo que insistimos es un ideal, pues la investigación siempre es realizada por seres humanos.

La validez, la confiabilidad y la objetividad no deben tratarse de forma separada. Sin alguna de las tres, el instrumento no es útil para llevar a cabo un estudio.

¿Cómo se sabe si un instrumento de medición es confiable y válido?

En la práctica es casi imposible que una medición sea perfecta y generalmente se tiene un grado de error. Desde luego, se trata de que este error sea el mínimo posible, por lo cual la medición de cualquier fenómeno se conceptualiza con la siguiente fórmula básica:

$$X = t + e$$

Donde X representa los valores observados (resultados disponibles); t, los valores verdaderos; y e, el grado de error en la medición. Si no hay error de medición (e es igual a cero), el valor observado y el verdadero son equivalentes. Esto puede verse claramente así:

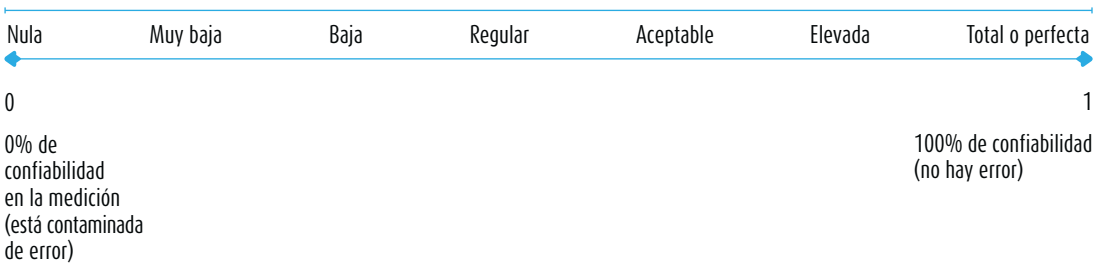
$$X = t + 0$$
$$X = t$$

Esta situación representa el ideal de la medición. Cuanto mayor sea el error al medir, el valor que observamos (en el cual nos basamos) se alejará más del valor real o verdadero. Por ejemplo, si medimos la motivación de un individuo y la medición está contaminada por un grado de error considerable, la motivación registrada por el instrumento será bastante diferente de la motivación real de esa persona. Por ello, es importante que el error se reduzca lo más posible. Pero, ¿cómo sabemos el grado de error que tenemos en una medición? Al calcular la confiabilidad y la validez.

Cálculo de la confiabilidad o fiabilidad

Hay diversos procedimientos para calcular la confiabilidad de un instrumento de medición. Todos utilizan procedimientos y fórmulas que producen coeficientes de fiabilidad. La mayoría oscilan entre cero y uno, donde un coeficiente de cero significa nula confiabilidad y uno representa un máximo de confiabilidad (fiabilidad total, perfecta). Cuanto más se acerque el coeficiente a cero, mayor error habrá en la medición. Esto se ilustra en la figura 9.4.

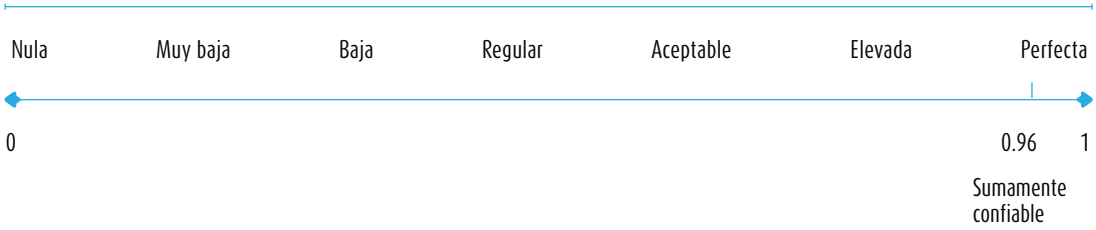
Figura 9.4 Interpretación de un coeficiente de confiabilidad.



Los procedimientos más utilizados para determinar la confiabilidad mediante un coeficiente son: 1) *medida de estabilidad* (confiabilidad por *test-retest*), 2) *método de formas alternativas o paralelas*, 3) *método de mitades partidas (split-halves)* y 4) *medidas de consistencia interna.* Estos procedimientos no se detallan en esta sección, sino que se explican en el capítulo 10, "Análisis de los datos cuantitativos", debido a que requieren conocimientos de ciertos conceptos estadísticos.[2] Simplemente se comentará su interpretación con la medida de congruencia interna denominada "coeficiente alfa Cronbach", que tal vez es la más utilizada.

En el caso de mediciones sociales, supongamos que una investigadora desarrolló un instrumento para medir el grado de "amor romántico" entre parejas de jóvenes universitarios, el cual se fundamentó en cuatro de las herramientas tal vez más conocidas para ello: la medida de Rubin sobre el amar y el vincularse con los demás, la escala sobre actitudes hacia el amor, la medida sobre el amor apasionado y la escala del amor triangular (Graham y Christiansen, 2009). Para estimar la confiabilidad de su instrumento lo debe aplicar a su muestra y sobre la base de los resultados calcular tal coeficiente. Imaginemos que obtiene un valor alfa Cronbach de 0.96, que es muy elevado, lo que significa que su medida del "amor romántico" es sumamente confiable, esto se representa en la figura 9.5.

Figura 9.5 Interpretación de un coeficiente de confiabilidad sobre un instrumento que mide el "amor romántico".



La confiabilidad varía de acuerdo con el número de indicadores específicos o ítems[3] que incluya el instrumento de medición. Cuantos más ítems haya, mayor tenderá a ser ésta, lo cual resulta lógico. Veámoslo con un ejemplo cotidiano: si se desea probar qué tan confiable o constante es la lealtad de un amigo hacia nuestra persona, cuantas más pruebas le pongamos, su fiabilidad será mayor. Claro está que demasiados ítems provocarán cansancio en los participantes.

Cada vez que se administra un instrumento de medición debe calcularse la confiabilidad, al igual que evaluarse la evidencia sobre la validez.

Cálculo de la validez

Con respecto a la validez de contenido, primero es necesario revisar cómo han medido la variable otros investigadores. Y, con base en dicha revisión, elaborar un universo de ítems o reactivos posibles para medir la variable y sus dimensiones (el universo debe ser lo más exhaustivo posible). Después, se consulta a investigadores familiarizados con la variable para ver si el universo es verdaderamente exhaustivo. Se seleccionan los ítems bajo una cuidadosa evaluación, uno por uno. Y si la variable está compuesta por diversas dimensiones o facetas, se extrae una muestra probabilística de reactivos, ya sea al azar o estratificada (cada dimensión constituiría un estrato). Se administran los ítems, se correlacionan las puntuaciones de éstos entre sí (tiene que haber correlaciones altas, en especial entre ítems que miden una misma dimensión, pero teniendo cuidado que sean capaces de discriminar entre participantes o casos) (Punch, 2009); y se hacen estimaciones estadísticas para ver si la muestra es representativa. Para calcular la validez de contenido son necesarios varios coeficientes. Éste sería un procedimiento

[2] Tales procedimientos se aplican a instrumentos o sistemas constituidos por escalas con varios reactivos, indicadores o mediciones (mediciones compuestas).

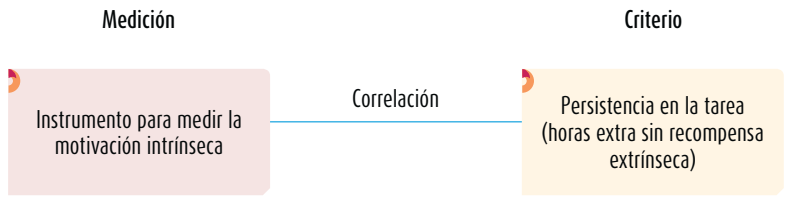
[3] Un ítem es la unidad mínima que compone una medición; es un reactivo que estimula una respuesta en un sujeto (por ejemplo, una pregunta, una frase, una lámina, una fotografía o un objeto de descripción).

www.elosopanda.com | jamespoetrodriguez.com

ideal. Pero, como veremos más adelante, a veces no se calculan estos coeficientes, sino que se seleccionan los ítems mediante un proceso que asegura la representatividad (no de manera estadística sino conceptual).

La *validez de criterio* se estima al correlacionar la medición con el criterio externo (puntuaciones del instrumento frente a las puntuaciones en el criterio), y este coeficiente se toma como coeficiente de validez (Bohrnstedt, 1976). Que podría representarse con el ejemplo de la figura 9.6.[4]

Figura 9.6 Ejemplo para la estimación de la validez de criterio.



La validez de constructo suele determinarse mediante procedimientos de análisis estadístico multivariado ("análisis de factores", "análisis discriminante", "regresiones múltiples", etc.), los cuales podrá bajar del centro de recursos en línea en: Material complementario → Capítulos → Capítulo 8: "Análisis estadístico: segunda parte".

Centro de recursos en línea

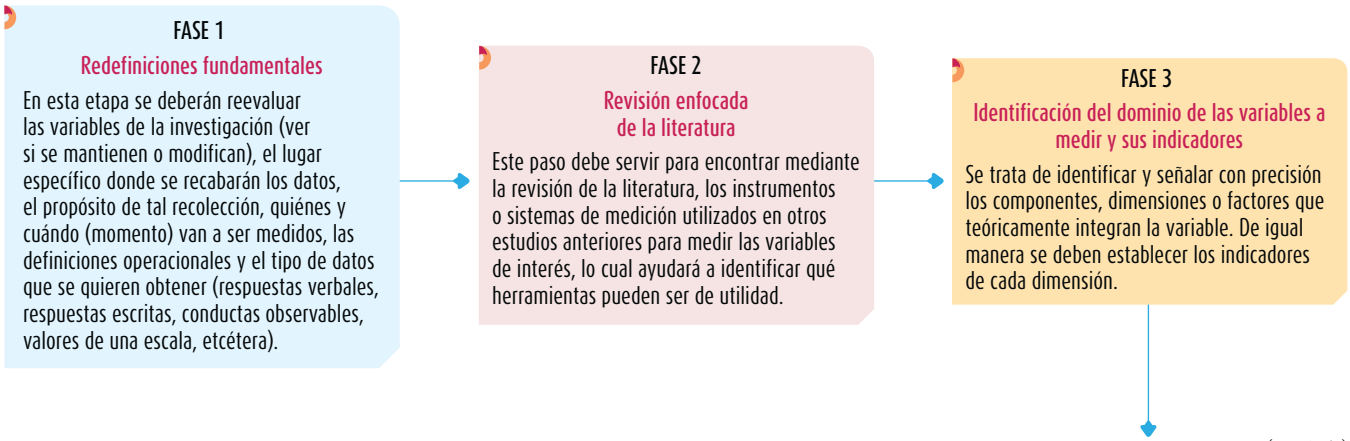
¿Qué procedimiento se sigue para construir un instrumento de medición?

5

Existen diversos tipos de instrumentos de medición, cada uno con características diferentes. Sin embargo, el procedimiento general para construirlos y aplicarlos es semejante. Éste se resume mediante etapas en el diagrama de la figura 9.7 y corresponde a la parte del plan de recolección que contesta a la pregunta: ¿a través de qué método vamos a recolectar los datos? Y cabe señalar que cada etapa o fase no se detalla en este capítulo, sino en el centro de recursos en línea en: Material complementario → Capítulos → Capítulo 7: "Recolección de los datos cuantitativos: segunda parte".

Centro de recursos en línea

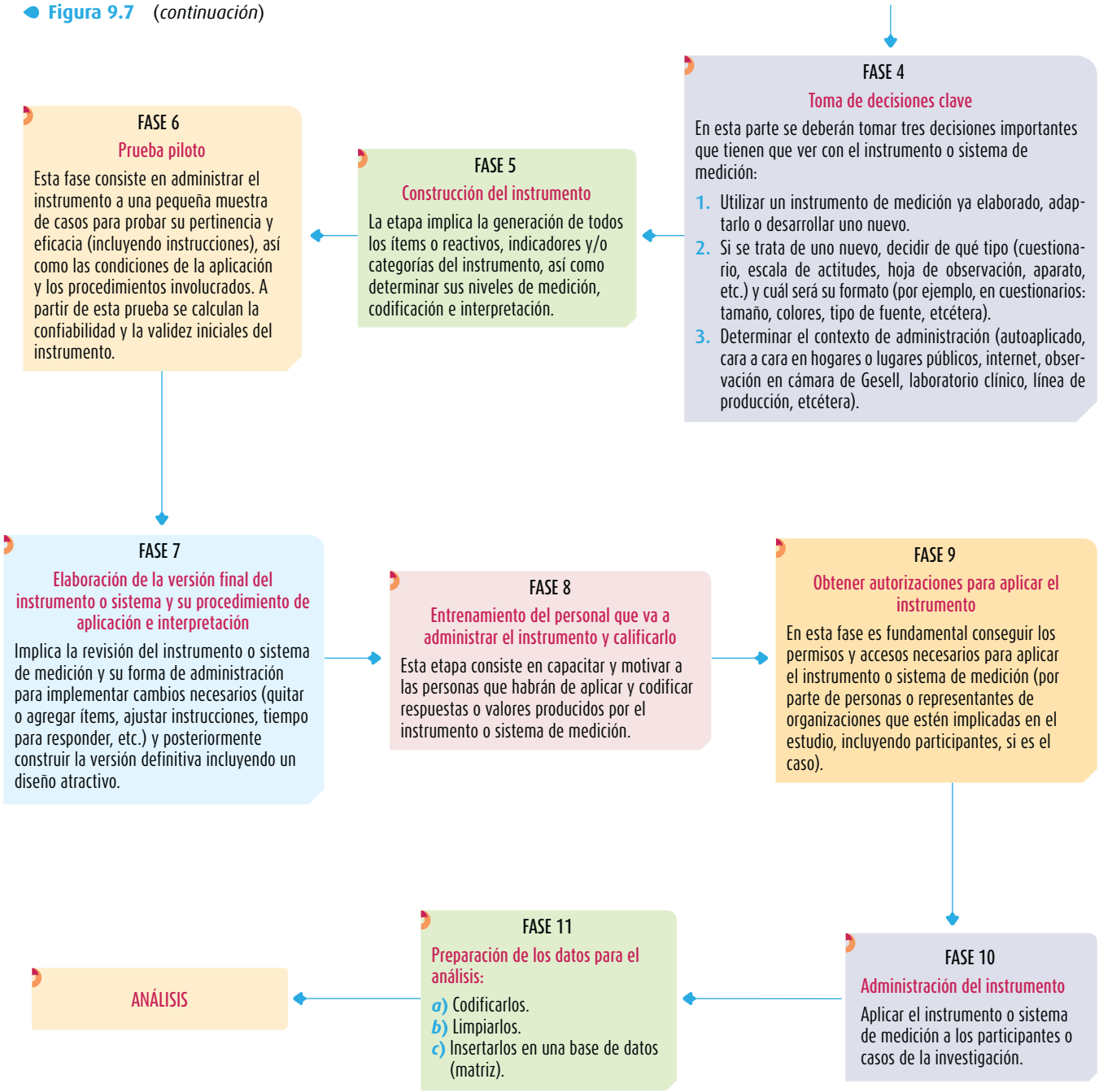
Figura 9.7 Proceso para construir un instrumento de medición.



(*continúa*)

[4] Las pruebas de correlación se presentan en el siguiente capítulo: "Análisis de los datos cuantitativos".

Figura 9.7 (*continuación*)



Las fases 1 a 7 del diagrama se refieren propiamente al desarrollo del instrumento o sistema de medición, mientras que las etapas 8 a 11 representan su administración y la preparación de los datos para su análisis.

Tres cuestiones fundamentales para un instrumento o sistema de medición

Existen tres cuestiones básicas respecto a las variables que deben considerarse al momento de construir un instrumento: *a*) la operacionalización, *b*) la codificación y *c*) establecer los niveles de medición.

La operacionalización: el tránsito de la variable al ítem o valor

El paso de una variable teórica a indicadores empíricos verificables y medibles e ítems o equivalentes se le denomina *operacionalización* (Solís, 2013). La operacionalización se fundamenta en la definición conceptual y operacional de la variable. Su proceso se amplía en el capítulo 7 adicional que se puede descargar del centro de recursos en línea. Por ahora, se comentará que cuando se construye un instrumento, el proceso más lógico para hacerlo es transitar de la variable a sus dimensiones o componentes, luego a los indicadores y finalmente a los ítems o reactivos y sus categorías. En la tabla 9.2 se dan ejemplos de este tránsito.

Tabla 9.2 Ejemplo de desarrollo de ítems

Estudio de las preferencias de los jóvenes para divertirse (ejemplo sencillo)[5]			
Variable	**Dimensión**	**Indicadores**	**Ítems**
Preferencia de actividad para salir con alguien del género opuesto	Actividad nocturna preferida entre semana	Jerarquía de preferencias de actividades de lunes a jueves (aunque algunos comienzan el fin de semana desde el jueves).	De lunes a jueves, ¿cuál sería tu actividad preferida nocturna para salir con el chico o chica que más te gusta? (marcar la que más te agrade). 1. Salir a cenar a un restaurante. 2. Ir al cine. 3. Ir a un bar, "antro", *grill*, taberna, *pub* o cervecería. 4. Acudir a una cafetería, heladería o a algún establecimiento en el que sirvan bebidas no alcohólicas. 5. Ir a bailar a una discoteca, disco o "antro". 6. Ir a una fiesta privada. 7. Acudir al teatro. 8. Acudir a un concierto. 9. Pasear por un parque, jardín o avenida. 10. Otra (especificar).
	Actividad nocturna preferida en fin de semana	Jerarquía de preferencias de actividades en viernes y sábado.	Mismas categorías u opciones de respuesta.
	Actividad nocturna preferida en domingo	Jerarquía de preferencias de actividades en domingo	Mismas categorías u opciones de respuesta.
Clima organizacional	Moral	Grado en que los miembros de una organización o departamento perciben que colaboran y cooperan entre sí, se apoyan mutuamente y mantienen relaciones de amistad y compañerismo.	• Mis compañeros de trabajo son mis amigos. 5. Totalmente de acuerdo. 4. De acuerdo. 3. Ni de acuerdo ni en desacuerdo. 2. En desacuerdo. 1. Totalmente en desacuerdo. • En mi trabajo hay mucho compañerismo (mismas opciones de respuesta que el ítem anterior). • Siempre que lo necesito mis compañeros de trabajo me brindan apoyo (mismas opciones de respuesta).

(*continúa*)

[5] En los ejemplos de la tabla sólo se incluyen unos cuantos ítems o reactivos por cuestiones de espacio, son ejemplos muy resumidos.

Tabla 9.2 (*continuación*)

Investigación sobre el clima organizacional			
Variable	**Dimensión**	**Indicadores**	**Ítems**
			• En el departamento donde trabajo nos mantenemos unidos (mismas opciones de respuesta. • La mayoría de las veces en mi departamento compartimos la información más que guardarla para nosotros (mismas opciones de respuesta). • ¿Qué tanto apoyo le brindan sus compañeros cuando usted lo necesita? 5. Total. 4. Bastante. 3. Aceptablemente. 2. Poco. 1. Ninguno.
	Autonomía	Grado de libertad percibida para tomar decisiones y realizar el trabajo.	• En esta empresa tengo libertad para tomar decisiones que tienen que ver con mi trabajo. 5. Totalmente de acuerdo. 4. De acuerdo. 3. Ni de acuerdo ni en desacuerdo. 2. En desacuerdo. 1. Totalmente en desacuerdo. • Mi jefe me da libertad para tomar decisiones que tienen que ver con mi trabajo (mismas opciones de respuesta que el ítem anterior.)
	Atribución del desempeño	Grado de conciencia compartida por desempeñarse con calidad en las tareas laborales, sobre la base de la cooperación.	• En esta empresa todos tratamos de hacer bien nuestro trabajo (mismas opciones de respuesta que el ítem anterior.) • En esta empresa todos queremos dar lo mejor de nosotros en el trabajo (mismas opciones.)

En ocasiones tenemos menos elementos, por ejemplo, sólo variable, dimensiones e ítems; incluso, únicamente variable y categorías de un único ítem o escala. Tal sería el caso del género —variable—, integrada por un ítem solitario —pregunta u observación— y dos categorías: masculino y femenino; o bien, temperatura —variable— y la lectura de un valor o categoría. Algunas variables son sencillas (un ítem) y otras compuestas (varias dimensiones, indicadores, ítems y categorías). En el siguiente capítulo "Análisis de los datos cuantitativos" (en el apartado "Apunte 1"), se discuten las implicaciones de lo anterior para el análisis estadístico. En el ejemplo de Lee y Guerin (2009) tratado en esta obra, la variable "satisfacción respecto a la calidad del diseño ambiental del interior del área de trabajo u oficina", incluyó la dimensión "satisfacción con respecto al diseño de la oficina", compuesta por tres indicadores: *a*) satisfacción sobre el espacio aprovechable para el trabajo individual y para guardar o almacenar, *b*) satisfacción en torno al nivel de privacidad visual y *c*) satisfacción en relación con la facilidad para interactuar con compañeros de trabajo; cada uno con uno o varios ítems. Ésta y otras variables fueron correlacionadas con la variable tipo de oficina, integrada por un ítem con las siguientes categorías:

1. Despacho u oficina privada cerrada.
2. Oficina cerrada compartida con otras personas.
3. Cubículo con paneles altos de metro y medio o más.
4. Cubículo con paneles bajos de menos de metro y medio.
5. Espacio de trabajo en una oficina abierta sin divisiones.
6. Otros.

Codificación

Codificar los datos significa asignarles un valor numérico o símbolo que los represente. Es decir, a las categorías (opciones de respuesta o valores) de cada ítem o variable se les asignan valores numéricos o signos que tienen un significado. Por ejemplo, si tuviéramos la variable "género" con sus respectivas categorías, masculino y femenino, a cada categoría le determinaríamos un valor. Esto podría ser:

Codificación Significa asignar a los datos un valor numérico o símbolo que los represente, ya que es necesario para analizarlos cuantitativamente.

Categoría	Codificación (valor asignado)
• Masculino	1
• Femenino	2

Por tanto, Paola Yáñez en la variable género sería 2, Luis Gerardo Vera y Ramón Calderón serían 1, Mabel González Bencosme 2 y así sucesivamente.

Otro ejemplo sería la variable "horas de exposición diaria a la televisión", que podría codificarse como se muestra en la tabla 9.3.

Tabla 9.3 Ejemplo de codificación

Categoría	Codificación (valor asignado)
— No ve televisión	0
— Menos de una hora	1
— Una hora	2
— Más de una hora, pero menos de dos	3
— Dos horas	4
— Más de dos horas, pero menos de tres	5
— Tres horas	6
— Más de tres horas, pero menos de cuatro	7
— Cuatro horas	8
— Más de cuatro horas	9

En el primer ejemplo de la tabla 9.2, la respuesta a la pregunta: *De lunes a jueves, ¿cuál sería tu actividad nocturna preferida para salir con el chico o chica que más te gusta?,* la codificación era con números (1 = salir a cenar a un restaurante; 2 = ir al cine; 3 = ir a un bar, "antro", grill, taberna, pub o cervecería; 4 = acudir a una cafetería, heladería o establecimiento en el que sirvan bebidas no alcohólicas; 5 = ir a bailar a una discoteca, disco o "antro"; 6 = ir a una fiesta privada; 7 = acudir al teatro; 8 = acudir a un concierto; 9 = pasear por un parque, jardín, avenida, y 10 = otra).

Mientras que en el ítem: "En esta empresa tengo libertad para tomar decisiones que tienen que ver con mi trabajo", la codificación era:

5. Totalmente de acuerdo
4. De acuerdo
3. Ni de acuerdo ni en desacuerdo
2. En desacuerdo
1. Totalmente en desacuerdo

Es necesario insistir que cada ítem o variable sencilla deberá tener una codificación (códigos numéricos o simbólicos) para sus categorías, a esto se le conoce como "precodificación". Desde luego, hay veces que un ítem no puede ser codificado *a priori* (precodificado), porque es muy difícil conocer cuáles serán sus categorías. Por ejemplo, si en una investigación preguntáramos: ¿qué opina del programa económico que recientemente aplicó el Gobierno? Las categorías podrían ser muchas más de

las que nos imaginemos y resultaría difícil predecir con precisión cuántas y cuáles serían. En tales situaciones, la codificación se lleva a cabo una vez que se aplica el ítem (*a posteriori*). Éste es el caso de algunos ítems que por ahora denominaremos "abiertos".

La codificación es necesaria para analizar cuantitativamente los datos (aplicar análisis estadístico). A veces se utilizan letras o símbolos en lugar de números (*, A, Z). La codificación puede o no incluirse en el instrumento de medición, veámoslo con un ejemplo de pregunta:

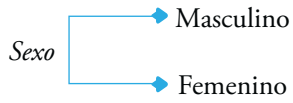
Pregunta precodificada ¿Tiene usted novia?		Pregunta no precodificada ¿Tiene usted novia?	
[1]	Sí	☐	Sí
[0]	No	☐	No

Asimismo, es muy importante indicar el nivel de medición de cada ítem y, por ende, el de las variables**,** porque es parte de la codificación y dependiendo de dicho nivel se selecciona uno u otro tipo de análisis estadístico (por ejemplo, la prueba estadística para correlacionar dos variables de intervalo es muy distinta de la prueba para correlacionar dos variables ordinales). De esta forma, es necesario hacer una relación de variables, ítems y niveles de medición.

Niveles de medición

Existen cuatro niveles de medición ampliamente conocidos.[6]

1. *Nivel de medición nominal.* En este nivel hay dos o más categorías del ítem o la variable. Las categorías no tienen orden ni jerarquía. Lo que se mide (objeto, persona, etc.) se coloca en una u otra categoría, lo cual indica tan sólo diferencias respecto de una o más características. Por ejemplo, la variable "género" de la persona posee sólo dos categorías: masculino y femenino. Ninguna de las categorías implica mayor jerarquía que la otra. Las categorías únicamente reflejan diferencias en la variable. No hay orden de mayor a menor.



Si les asignamos una etiqueta o un símbolo a cada categoría, esto identificará exclusivamente a la categoría. Por ejemplo:

$$* = \text{Masculino}$$
$$z = \text{Femenino}$$

Si usamos numerales, es lo mismo:

1 = Masculino		**2** = Masculino
	es igual a	
2 = Femenino		**1** = Femenino

Los números utilizados en este nivel de medición tienen una función puramente de clasificación *y no* se pueden manipular de manera aritmética. Por ejemplo, la afiliación religiosa es una variable nominal; si pretendiéramos operarla de forma aritmética se presentarían situaciones tan ridículas como ésta:

[6] Para ampliar este tema se sugiere Issel (2014), Wolf (2008), Stoutenborough (2008), Dykema, Blixt y Stevenson (2008); Scanlan (2008), Gershkoff (2008) y Rasmussen (2006).

1 = Católico	
2 = Judío	**1 + 2 = 3**
3 = Protestante	
4 = Musulmán	¿Un católico + un judío = un protestante?
5 = Otros	(No tiene sentido)

Las variables nominales pueden incluir dos categorías (dicotómicas), o bien, tres o más categorías (categóricas). Ejemplos de variables nominales dicotómicas serían el género, el veredicto de un jurado (culpable-no culpable) y el tipo de escuela a la que se asiste (privada-pública); y como ejemplos de variables nominales categóricas tendríamos la afiliación política (partido A, partido B, etc.), la licenciatura estudiada, el grupo étnico, el departamento, la provincia o el estado de nacimiento, la clase de material de construcción ("no" su resistencia, que sería otra variable), tipo de medicamento administrado ("no" la dosis, que sería una variable distinta), bloques de mercado (asiático, latinoamericano, comunidad europea, etc.) y el canal de televisión preferido.

2. *Nivel de medición ordinal*. En este nivel hay varias categorías, pero además mantienen un orden de mayor a menor. Las etiquetas o los símbolos de las categorías sí indican jerarquía. Por ejemplo, el prestigio ocupacional en Estados Unidos se ha medido por diversas escalas que reordenan las profesiones de acuerdo con su prestigio, por ejemplo:[7]

Valor en escala	**Profesión**
90	Ingeniero químico
80	Científico de ciencias naturales (excluyendo la química)
60	Actor común
50	Operador de estaciones eléctricas de potencia
02	Manufacturero de tabaco

Los números (símbolos de categorías) definen posiciones, en el ejemplo: 90 es más que 80, 80 más que 60, 60 más que 50 y así sucesivamente. Sin embargo, las categorías no están ubicadas a intervalos iguales (no hay un intervalo común). No podríamos decir con exactitud que entre un actor (60) y un operador de estaciones eléctricas (50) existe la misma distancia en prestigio que entre un científico de ciencias naturales (80) y un ingeniero químico (90). Al parecer, en ambos casos la distancia es 10, pero no es una distancia real. Otra escala[8] clasificó el prestigio de dichas profesiones de la siguiente manera:

Valor en escala	**Profesión**
98	Ingeniero químico
95	Científico de ciencias naturales (excluyendo la química)
84	Actor común
78	Operador de estaciones eléctricas de potencia
13	Manufacturero de tabaco

Aquí la distancia entre un actor (84) y un operador de estaciones (78) es de seis, y la distancia entre un ingeniero químico (98) y un científico de ciencias naturales (95) es de tres. Otro ejemplo sería la variable *posición jerárquica en la empresa*:

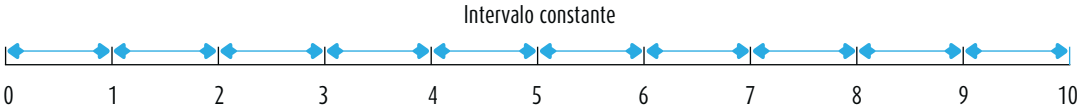
Presidente	10	Jefe	5
Vicepresidente	9	Empleado *A*	4
Director general	8	Empleado *B*	3
Gerente de área	7	Empleado *C*	2
Subgerente o superintendente	6	Intendencia	1

[7] Duncan (1977). La escala incluye múltiples profesiones, se presentan sólo algunas.
[8] Nam *et al.* (1965) y Nam (1983). Para escalas más complejas y actuales, véase Organización Internacional del Trabajo (2012).

Sabemos que el presidente (10) es más que el vicepresidente (9); éste, más que el director general (8); a su vez este último, más que el gerente (7) y así sucesivamente; pero no se precisa en cada caso cuánto más. Tampoco pueden utilizarse las operaciones aritméticas básicas: no podríamos decir que 4 (empleado *A*) + 5 (jefe) = 9 (vicepresidente), ni que 10 (presidente) ÷ 5 (jefe) = 2 (empleado C). Sería absurdo, no tiene sentido. Otros ejemplos de este nivel serían: la medición por rangos de las preferencias de marcas de bebidas refrescantes con gas (refrescos o sodas), autopercepción del grado de dolor de cabeza, percepción de sabores respecto a un platillo y jerarquización de valores (en primer lugar, en segundo lugar, en tercero).

3. *Nivel de medición por intervalos.* Además del orden o la jerarquía entre categorías, se establecen intervalos iguales en la medición. Las distancias entre categorías son las mismas a lo largo de toda la escala, por lo que hay un intervalo constante, una unidad de medida.

Figura 9.8 Escala con intervalos iguales entre categorías.



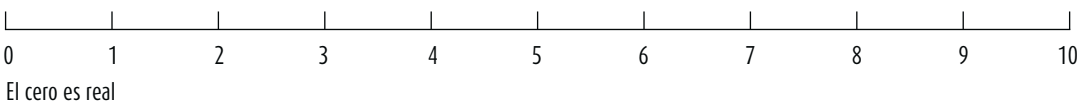
Por ejemplo, en una prueba de resolución de problemas matemáticos (30 problemas de igual dificultad). Si Ana Cecilia resolvió 10, Laura resolvió 20 y Abigail, 30. La distancia entre Ana Cecilia y Laura es igual a la distancia entre Laura y Abigail.

Sin embargo, en este nivel de medición el cero (0) es arbitrario, no es real, ya que se asigna arbitrariamente a una categoría el valor de cero y a partir de ésta se construye la escala. Un ejemplo clásico en ciencias naturales es la temperatura, que puede medirse en grados centígrados y Fahrenheit: el cero es arbitrario, pues no implica que en realidad haya cero (ninguna) temperatura (incluso en ambas escalas el cero es diferente).

Cabe agregar que diversas mediciones en el estudio del comportamiento humano no son verdaderamente de intervalo (por ejemplo, escalas de actitudes, pruebas de inteligencia y de otros tipos); pero para algunos autores se acercan a este nivel y se suele tratarlas como si fueran mediciones de intervalo.[9] Esto se hace porque este nivel de medición permite utilizar las operaciones aritméticas básicas y diversos métodos estadísticos, que de otro modo no se utilizarían. Aunque algunos investigadores no están de acuerdo con suponer tales mediciones como si fueran de intervalo. El producto interno bruto o producto nacional bruto estaría en este estadio.

4. *Nivel de medición de razón.* En este nivel, además de tenerse todas las características del nivel de intervalos (unidad de medida común, intervalos iguales entre las categorías y aplicación de operaciones aritméticas básicas y sus derivaciones), el cero es real y absoluto (no es arbitrario). Cero absoluto implica que hay un punto en la escala donde está ausente o no existe la propiedad medida (véase la figura 9.9).

Figura 9.9 Ejemplo de escala para el nivel de medición de razón.



Ejemplos de estas mediciones serían la exposición a la televisión (en minutos), el número de hijos, las ventas de un producto, los metros cuadrados de construcción, escala de temperatura de Kelvin, ingresos en una unidad monetaria (alguien puede tener cero ingresos), consumo de un medicamento específico, etcétera.

[9] Este comentario nos remite a una polémica que se comentará más adelante en el tema de la escala de Likert.

www.elosopanda.com | jamespoetrodriguez.com

¿De qué tipos de instrumentos de medición o recolección de datos cuantitativos disponemos en la investigación?

4

En la investigación disponemos de múltiples tipos de instrumentos para medir las variables de interés y en algunos casos llegan a combinarse varias técnicas de recolección de los datos. A continuación describimos brevemente las más usuales.

Los instrumentos que serán revisados en este capítulo son: cuestionarios[10] y escalas de actitudes, mencionando a otros. En el capítulo adicional 7, "Recolección de los datos cuantitativos, segunda parte", que se puede descargar en el centro de recursos en línea, se comentan los siguientes: registros del contenido (análisis de contenido) y observación cuantitativa, pruebas estandarizadas, recolección de información factual e indicadores (análisis de datos secundarios de registros públicos y documentación) y metaanálisis, así como otras clases de mediciones. Los metodólogos de cada ciencia y disciplina han desarrollado una gran diversidad de instrumentos y procedimientos, algunos de los cuales puede buscarlos el lector en los manuales y ejemplos del centro de recursos (http://www.mhhe.com/he/hmi6e).

La codificación y la preparación de los datos obtenidos se discutirán después de presentar los principales instrumentos de medición.

Cuestionarios

4

En fenómenos sociales, tal vez el instrumento más utilizado para recolectar los datos es el cuestionario. Un **cuestionario** consiste en un conjunto de preguntas respecto de una o más variables a medir (Chasteauneuf, 2009). Debe ser congruente con el planteamiento del problema e hipótesis (Brace, 2013).

Cuestionario Conjunto de preguntas respecto de una o más variables que se van a medir.

Los cuestionarios se utilizan en encuestas de todo tipo (por ejemplo, para calificar el desempeño de un gobierno, conocer las necesidades de hábitat de futuros compradores de viviendas y evaluar la percepción ciudadana sobre ciertos problemas como la inseguridad). Pero también, se implementan en otros campos. Por ejemplo, un ingeniero en minas usó un cuestionario como herramienta para que expertos de diversas partes del mundo aportaran opiniones calificadas con el fin de resolver ciertas problemáticas de producción.

Ahora, comentaremos primero sobre las preguntas y luego sobre las características deseables de este tipo de instrumento, así como los contextos en los cuales se pueden administrar los cuestionarios.

¿Qué tipos de preguntas se pueden elaborar?

5

El contenido de las preguntas de un cuestionario es tan variado como los aspectos que mide. Básicamente se consideran *dos tipos de preguntas: cerradas y abiertas.*

Preguntas cerradas

Las **preguntas cerradas** contienen categorías u opciones de respuesta que han sido previamente delimitadas. Es decir, se presentan las posibilidades de respuesta a los participantes, quienes deben acotarse a éstas. Pueden ser dicotómicas (dos posibilidades de respuesta) o incluir varias opciones de respuesta. Ejemplos de preguntas cerradas dicotómicas serían:

Preguntas cerradas Son aquellas que contienen opciones de respuesta previamente delimitadas. Resultan más fáciles de codificar y analizar.

¿Estudia usted actualmente?

( ) Sí
( ) No

¿Durante la semana pasada vio la final de la Liga de Campeones de Europa?

( ) Sí
( ) No

[10] Las entrevistas se plantean como un contexto en el cual pueden ser administrados los cuestionarios.

Ejemplos de preguntas cerradas con varias opciones de respuesta serían:

Como usted sabe, todos los países desarrollados reciben inmigrantes. ¿Cree que, en términos generales, la inmigración es más bien positiva o más bien negativa para estos países?

- ○ Positiva
- ○ Ni positiva ni negativa
- ○ Negativa
- ○ No sabría decir

¿Cuál es el puesto que ocupa usted en su empresa?

- ○ Presidente/Director general
- ○ Subdirector/Director/Gerente
- ○ Coordinador
- ○ Supervisor
- ○ Obrero
- ○ Vicepresidente/Director corporativo
- ○ Subgerente/Superintendente
- ○ Jefe de área
- ○ Empleado
- ○ Otro (especificar) _____

Si usted tuviera elección, ¿preferiría que su salario fuera de acuerdo con su productividad en el trabajo?

- ○ Definitivamente sí
- ○ Probablemente sí
- ○ No estoy seguro
- ○ Probablemente no
- ○ Definitivamente no

Como se observa, en las preguntas cerradas las categorías de respuesta son definidas *a priori* por el investigador y se le muestran al encuestado, quien debe elegir la opción que describa más adecuadamente su respuesta. Gambara (2002) hace notar algo muy lógico pero que en ocasiones se descuida y resulta fundamental: cuando las preguntas presentan varias opciones, éstas deben recoger todas las posibles respuestas.

Ahora bien, hay preguntas cerradas en las que el participante puede seleccionar más de una opción o categoría de respuesta (*posible multirrespuesta*).

Ejemplo

Supongamos que un entrevistador pregunta:

¿Esta familia tiene en el hogar...? (*Marque con una cruz o tache todas las opciones que el entrevistado o entrevistada señale que tiene en su hogar*):

- ○ Radio
- ○ Reproductor de DVD o Blu-ray
- ○ Computadora fija
- ○ Tableta, laptop o iPad
- ○ Teléfono celular o móvil
- ○ iPod u otro dispositivo similar
- ○ Televisión
- ○ TV de paga (SKY, Cablevisión, DirectTV, otros sistemas locales de cable o TV satelital)
- ○ Internet
- ○ Equipo de sonido para CD
- ○ Teléfono (línea telefónica en casa)

En preguntas como la del ejemplo anterior, los participantes pueden marcar una, dos, tres, cuatro o más opciones de respuesta. Las categorías no son mutuamente excluyentes.

En ocasiones, el encuestado tiene que jerarquizar opciones.

Ejemplo

De las siguientes compañeras de clase, ¿quién te atrae más? ¿Cuál en segundo lugar? ¿Cuál en tercer lugar? ¿Cuál en cuarto lugar? ¿Cuál en último lugar?

- ○ Sandra
- ○ Lucía
- ○ Ana
- ○ Mariana
- ○ Paola

O bien, en otras preguntas se debe designar un puntaje a una o diversas cuestiones.

Ejemplo

A continuación voy a mencionarle algunos de los problemas que suelen preocupar a los habitantes de este municipio y le pediría que en cada caso me dijera: ¿qué tanto le preocupa a usted cada uno de ellos?; donde 10 significa: "me preocupa muchísimo" y 0 quiere decir: "no me preocupa en absoluto".[11]

___ Desempleo
___ Pobreza
___ Inseguridad al transitar por la calle o viajar en transporte público
___ Empleo mal remunerado/bajos salarios
___ Robos/asaltos en los hogares y viviendas
___ Robos de vehículos/autos, motocicletas, bicicletas
___ Pandillerismo
___ Venta de drogas-narcomenudeo
___ Secuestros
___ Recolección de la basura (no todos los días la recogen)
___ Horarios inadecuados para la recolección de la basura
___ Escasez de agua
___ Cortes en el suministro de agua
___ Falta de vivienda
___ Servicios de salud insuficientes
___ Carencia/deficiencia de servicios educativos
___ Drenaje inadecuado en las calles
___ Tránsito/tráfico/vialidad
___ Pavimentación y bacheo mal hechos
___ Falta de infraestructura (calles, puentes, etcétera)
___ Corrupción de funcionarios municipales, policías, agentes de tránsito y vialidad
___ Situación económica familiar

En otras preguntas, se anota una cifra dentro de un rango predeterminado.

Ejemplo

Aproximadamente, ¿cuántos minutos dedica diariamente a hacer deporte entre semana, es decir, de lunes a viernes?

En algunas más, el encuestado se ubica en una escala. El concepto de escala (aplicado a la medición) puede definirse como: "sucesión ordenada de valores distintos de una misma cualidad" (Real Academia Española, 2001, p. 949). Es un patrón, conjunto, medida o estimación regular de acuerdo con algún estándar o tasa, respecto de una variable (Carmines y Woods, 2003b). Ejemplos: escala de temperatura en grados centígrados, escala de inteligencia, escala de distancia en kilómetros, metros y centímetros; escala de peso en kilogramos, escala de presión arterial en milímetros de mercurio (mmHg), escala musical con octavas, etcétera.

Ejemplo

¿Qué tan enamorada está usted de su novio? (del 0 al 100):

☺	100 – Completamente enamorada		.
	99		.
	98		.
	.		20
	.		10
	.		.
	80		.
	70		2
	60		1
☹	50	☹	0 – Nada enamorada

[11] Esta pregunta fue administrada a personas que tuvieran estudios mínimos de bachillerato o preparatoria, aunque funcionó con personas cuyo nivel era de secundaria. No se incluyeron todos los problemas por cuestiones de espacio, solamente algunos para ilustrar el tipo de pregunta.

Finalmente, en ocasiones se encadenan varias preguntas en una, como en el siguiente ejemplo (los candidatos y candidatas son ficticios).

Ejemplo

Le voy a mencionar algunos nombres de políticos de nuestro municipio y le pediría que en cada caso me dijera si sabe usted quién es y a qué partido pertenece, así como su opinión de esta persona.

Político (Rotar opciones)	P. 8 ¿Sabe quién es?		*Cuando sabe quién es* P. 9 ¿Sabe a qué partido pertenece? (*No leer opciones*)		*Cuando sabe quién es* P. 10 ¿Qué tan favorable o desfavorable es su opinión acerca de...? (*Leer opciones*)			
	Sí	*No (Pasar a p. 16)*	*Sí identificó*	*No identificó*	*Muy favorable*	*Favorable*	*Desfavorable*	*Muy desfavorable*
Guadalupe Méndez Peña	1	2	(Partido 1) 1	2	4	3	2	1
Agustín Almanza Mendoza	1	2	(Partido 2) 1	2	4	3	2	1
Sandra Hernández Jiménez	1	2	(Partido 3) 1	2	4	3	2	1
Roberto Yáñez Ruiz	1	2	(Partido 4) 1	2	4	3	2	1

Preguntas abiertas No delimitan las alternativas de respuesta. Son útiles cuando no hay suficiente información sobre las posibles respuestas de las personas.

Preguntas abiertas

En cambio, las **preguntas abiertas** no delimitan de antemano las alternativas de respuesta, por lo cual el número de categorías de respuesta es muy elevado; en teoría, es infinito, y puede variar de población en población.

Ejemplo

¿Por qué asiste a psicoterapia?

_____

_____

_____

¿Qué opina de las medidas de apoyo a la población que adoptó el Gobierno para disminuir el impacto del último terremoto ocurrido el 10 de noviembre?

_____

_____

_____

¿Conviene usar preguntas cerradas o abiertas?

Un cuestionario obedece a diferentes necesidades y a un problema de investigación, lo cual origina que en cada estudio el tipo de preguntas sea distinto. Algunas veces se incluyen tan sólo preguntas cerradas, otras ocasiones únicamente preguntas abiertas, y en ciertos casos ambos tipos de preguntas. Cada clase de pregunta tiene sus ventajas y desventajas, las cuales se mencionan a continuación.

Las *preguntas cerradas* son más fáciles de codificar y preparar para su análisis. Asimismo, estas preguntas requieren un menor esfuerzo por parte de los encuestados, que no tienen que escribir o verbalizar pensamientos, sino únicamente seleccionar la alternativa que sintetice mejor su respuesta. Responder a un cuestionario con preguntas cerradas toma menos tiempo que contestar uno con preguntas abiertas. Cuando el cuestionario se envía por correo, se tiene un mayor grado de respuesta porque es fácil de contestar y completarlo requiere menos tiempo. Otras ventajas son: se reduce la ambigüedad de las respuestas y se favorecen las comparaciones entre las respuestas (Burnett, 2009).

La *principal desventaja de las preguntas cerradas* reside en que limitan las respuestas de la muestra y, en ocasiones, ninguna de las categorías describe con exactitud lo que las personas tienen en mente; no siempre se captura todo "lo que pasa por la cabeza" de los participantes. Su redacción exige mayor laboriosidad (Vinuesa, 2005).

Para *formular preguntas cerradas* es necesario anticipar las posibles alternativas de respuesta. De no ser así, es muy difícil plantearlas. Además, el investigador debe asegurarse de que los participantes a quienes se les administrarán conocen y comprenden las categorías de respuesta. Por ejemplo, si preguntamos qué canal de televisión es el preferido, determinar las opciones de respuesta y que los participantes las comprendan es muy sencillo. Pero si preguntamos sobre las razones que provocan esa preferencia, señalar las alternativas es algo más complejo.

Las *preguntas abiertas* proporcionan una información más amplia y son particularmente útiles cuando no tenemos información sobre las posibles respuestas de las personas o la que tenemos es insuficiente (Phillips, Phillips y Aaron, 2013). También sirven en situaciones donde se desea profundizar una opinión o los motivos de un comportamiento. Su mayor desventaja es que son más difíciles de codificar, clasificar y preparar para el análisis. Además, llegan a presentarse sesgos derivados de distintas fuentes; por ejemplo, quienes enfrentan dificultades para expresarse en forma oral y por escrito quizá no respondan con precisión a lo que en realidad desean, o generen confusión en sus respuestas. El nivel educativo, la capacidad de manejo del lenguaje y otros factores pueden afectar la calidad de las respuestas (Saris y Gallhofer, 2007 y Black y Champion, 1976). Asimismo, responder a preguntas abiertas requiere mayor esfuerzo y más tiempo.

La elección del tipo de preguntas que contenga el cuestionario depende del grado en que se puedan anticipar las posibles respuestas, los tiempos de que se disponga para codificar y si se quiere una respuesta más precisa o profundizar en alguna cuestión. Una recomendación para construir un cuestionario es que se analice, variable por variable, qué tipo de pregunta o preguntas suelen ser más confiables y válidas para medir esa variable, de acuerdo con la situación del estudio (planteamiento del problema, características de la muestra, tipo de análisis a efectuar, etcétera).

Con frecuencia, las preguntas cerradas se construyen con fundamento en preguntas abiertas. Por ejemplo, en la prueba piloto puede elaborarse una pregunta abierta y posteriormente a su aplicación, sobre la base de las respuestas, se genera el ítem cerrado.

Figura 9.10 Ejemplo del paso de una pregunta abierta a la elaboración de una pregunta cerrada.

¿Una o varias preguntas para medir una variable?

En ocasiones sólo basta una pregunta para recolectar la información necesaria sobre la variable considerada, no es necesario incluir más. Por ejemplo, para medir el nivel de escolaridad de una muestra basta con preguntar: ¿hasta qué año escolar cursó?, o ¿cuál es su grado máximo de estudios? En otras

ocasiones se requiere elaborar varias preguntas para verificar la consistencia de las respuestas u obtener múltiples indicadores de diversas dimensiones de la variable medida.

Por ejemplo, algunas asociaciones latinoamericanas de investigación de mercados e instituciones educativas miden el nivel socioeconómico tomando en cuenta diversas preguntas o indicadores de la clase social, el estilo de vida, el ingreso y el bienestar:[12]

1. Escolaridad del jefe del hogar.
2. Número de focos en la vivienda.
3. Número de habitaciones en la vivienda, sin incluir baños.
4. Número de baños en general y con regadera.
5. Número de automóviles y otros vehículos en la cochera.
6. Posesión de ciertos aparatos y equipos (tanque de gas estacionario, boiler o calentador, estufa de gas, lavadora automática, aspiradora, horno de microondas, tostadora de pan, etcétera).
7. Posesión de elementos de conectividad y entretenimiento (computadora, televisión, equipo de sonido, equipo de video, línea telefónica, teléfono móvil o celular, internet, servicio de televisión de paga, etcétera).
8. Características de la vivienda (techo estable y seguro en su vivienda —no de cartón ni hule, etc.—, piso firme en su interior —cemento, concreto o piso de mosaico—, agua que llega por tubería a su vivienda, etcétera).
9. Tarjetas de crédito comercial en el hogar.

Con base en estos indicadores se construyen índices, cada uno con un peso o coeficiente, y al final se otorga una puntuación que determina el nivel socioeconómico con mayor precisión.[13] Sin embargo, esto puede resultar muy complejo para el alumno que comienza con sus primeras investigaciones, por lo cual la alternativa sería preguntar a los miembros de la familia que trabajan aproximadamente cuál es su nivel mensual de ingresos y cuántos focos eléctricos tiene aproximadamente en su casa.[14]

Así, es recomendable *hacer solamente las preguntas necesarias* para obtener la información deseada o medir la variable.

Una modalidad de cuestionamientos múltiples lo es la batería de preguntas, la cual sirve para: *a*) ahorrar espacio en el cuestionario, *b*) facilitar la comprensión del mecanismo de respuesta (si se entiende la primera pregunta, se comprenderán las demás) (Corbetta, 2003) y *c*) construir índices que permitan obtener una calificación total.

Ejemplo

Variable a medir: visión departamental

Definición conceptual: percepción de la meta departamental en cuanto a claridad, naturaleza visionaria, grado en que es posible alcanzarla y medida en que puede ser compartida, y que representa una fuerza motivacional para el trabajo (Anderson y West, 1998 y Hernández-Sampieri, 2005).[15]

[12] Asociación Mexicana de Agencias de Investigación de Mercados y Opinión Pública, AMAI (2013), López (2009) y Universidad de Celaya (2013). Con el fin de conocer el método por puntos para la ubicación del nivel socioeconómico, se recomienda consultar directamente estas fuentes o consultar a la asociación de empresas de investigación de mercados de su país.

[13] Desde luego, estos índices evolucionan permanentemente, en el corto plazo se contempla agregar indicadores respecto al ocio y tiempo libre, salud y esperanza de vida, idiomas, información y conocimiento y balance de vida (López, 2011).

[14] En varios estudios se ha demostrado que el nivel de ingresos está relacionado con el número de focos de una casa habitación (residencia, hogar o propiedad). El número de focos se vincula con el número de cuartos, extensión de la casa, tamaño del jardín y otros factores, es decir, con el valor de la propiedad (Universidad de Celaya, 2013). Los rangos podrían ser: 3 focos o menos: estratos muy desfavorecidos; de 4 a 5 focos: estratos desfavorecidos; de 6 a 10 focos: estratos medios característicos; 11 a 15: estratos medios favorecidos; 16 a 20: medios/altos; 21 a 30: altos favorecidos; más de 31: muy altos o completamente favorecidos. Se sugiere excluir candiles. En cada nación cambia la designación de cada estrato, y no se quiere utilizar los términos "bajos", que son peyorativos.

[15] El ejemplo contiene solamente algunas de las preguntas de la escala original. Asimismo, se conjuntaron con otras escalas y se distribuyeron a lo largo del cuestionario.

Preguntas o ítems:	Completamente (mucho)	Aceptablemente	Regular	Poco	Nada
1. ¿Qué tan claros tiene los objetivos de su departamento?	5	4	3	2	1
2. ¿En qué medida considera usted que los objetivos de su departamento son útiles y apropiados?	5	4	3	2	1
3. ¿Qué tan de acuerdo está usted con estos objetivos de su departamento?	5	4	3	2	1
4. ¿En qué medida piensa usted que los objetivos de su departamento son claros?	5	4	3	2	1
5. ¿En qué medida piensa usted que los objetivos de su departamento son comprendidos por sus compañeros de trabajo del mismo departamento?	5	4	3	2	1
6. ¿En qué medida considera usted que sus compañeros de departamento están de acuerdo con los objetivos?	5	4	3	2	1
7. En qué medida considera que los objetivos del departamento pueden lograrse actualmente?	5	4	3	2	1

¿Las preguntas van precodificadas o no?

Siempre que se pretenda efectuar análisis estadístico, se requiere codificar las respuestas de los participantes a las preguntas del cuestionario, y debemos recordar que esto significa asignarles símbolos o valores numéricos y que cuando se tienen preguntas cerradas es posible codificar *a priori* o *precodificar* las opciones de respuesta, e incluir esta precodificación en el cuestionario (como en el último ejemplo).

Ejemplo

De preguntas precodificadas

¿Tiene usted inversiones en la Bolsa de Valores?

☐ 1 Sí ☐ 2 No

Cuando usted se enfrenta a un problema en su trabajo, para resolverlo recurre generalmente a:

1. Su superior inmediato
2. Su propia experiencia
3. Sus compañeros
4. Los manuales de políticas y procedimientos
5. Otra fuente: _____
(especificar)

En ambas preguntas, las respuestas van acompañadas de su valor numérico correspondiente, es decir, se han precodificado. Obviamente en las preguntas abiertas la codificación se realiza después, una vez que se tienen las respuestas. Las preguntas y opciones de respuesta precodificadas poseen la ventaja de que su codificación y preparación para el análisis son más sencillas y requieren menos tiempo.

¿Qué preguntas son obligatorias?

Regularmente, las preguntas llamadas demográficas o de ubicación del participante encuestado: género, edad, nivel socioeconómico, estado civil, escolaridad (nivel de estudios), religión, afiliación política, colonia, barrio o zona donde vive, pertenencia a ciertas agrupaciones, ocupación (actividad a la que se dedica), años de vivir en el lugar actual de residencia, etc. En empresas: puesto, antigüedad,

área funcional donde trabaja (gerencia, departamento, dirección o equivalente), planta u oficinas donde labora y demás preguntas. En cada investigación debemos analizar cuáles son pertinentes y nos resultarán útiles.

¿Qué características debe tener una pregunta?

Independientemente de que las preguntas sean abiertas o cerradas, y de que sus respuestas estén precodificadas o no, hay una serie de características que deben cubrirse al plantearlas:

a) Las preguntas tienen que ser claras, precisas y comprensibles para los sujetos encuestados. Deben evitarse términos confusos, ambiguos y de doble sentido. Por ejemplo, la pregunta: "¿ve usted televisión?", es confusa, no delimita cada cuánto. Sería mucho mejor especificar: ¿acostumbra usted ver televisión diariamente?, ¿cuántos días durante la última semana vio televisión?, y después preguntar horarios, canales y contenidos de los programas. Otro ejemplo inconveniente sería: ¿le gusta el deporte? No se sabe si se trata de verlo por televisión o en vivo, si de practicarlo o qué, y en última instancia, ¿cuál deporte? Otro caso que genera confusión son los términos con múltiples significados (Burnett, 2009), por ejemplo: ¿su empleo es estable?, implica un concepto de estabilidad de empleo que no tiene un solo significado. ¿Qué se considera estable?: ¿un contrato por un año, por dos, por cinco...?

Un caso común de confusión son las palabras sobre la temporalidad, resulta nebuloso el cuestionamiento: ¿ha asistido recientemente al cine?, ya que implica otras preguntas: ¿qué significa recientemente?, ¿ayer, la última semana, el último mes? Sería mejor interrogar: durante las últimas dos semanas (o mes), ¿cuántas veces ha ido al cine? De igual forma: ¿ha trabajado desde joven?, habrá de sustituirse por: ¿a partir de qué edad comenzó a trabajar?

b) Es aconsejable que las preguntas sean lo más breves posible, porque las preguntas largas suelen resultar tediosas, toman más tiempo y pueden distraer al participante; pero como menciona Rojas (2002) no es recomendable sacrificar claridad por concisión. Cuando se trata de asuntos complicados tal vez es mejor una pregunta más larga, debido a que facilita el recuerdo, proporciona al sujeto más tiempo para pensar y favorece una respuesta más articulada (Corbetta, 2003). La directriz es que se incluyan las palabras necesarias para que se comprenda la pregunta, sin ser repetitivos o barrocos.

c) Deben formularse con un vocabulario simple, directo y familiar para los participantes.

d) No pueden incomodar a la persona encuestada ni ser percibidas como amenazantes y nunca ésta debe sentir que se le enjuicia. Debemos inquirir de manera sutil. Preguntas como: ¿acostumbra consumir algún tipo de bebida alcohólica?, tienden a provocar rechazo. Es mejor cuestionar: ¿algunos de sus amigos acostumbran consumir cierto tipo de bebida alcohólica?, y después utilizar preguntas tenues que indirectamente nos indiquen si la persona acostumbra consumir esta clase de bebidas (¿cuál es su tipo de bebida favorita?, ¿cada cuánto se reúne con sus amigos?, etc.). Mertens (2010) sugiere sustituir la pregunta: ¿es usted alcohólico? (en extremo amenazante), por la siguiente formulación: "el consumo de bebidas como el ron, tequila, vodka y whisky en esta ciudad es de *X* botellas de un litro, ¿en qué medida usted estaría por encima o por debajo de esta cantidad?" (alternativas de respuesta: "por encima, igual o por debajo"). Gochros (2005) recomienda cambiar la pregunta: ¿consume drogas?, por: ¿qué opina de las personas que consumen drogas en dosis mínimas? En estos casos de preguntas difíciles es posible usar escalas de actitud en lugar de preguntas o aun otras formas de medición (como se verá en la parte de escalas actitudinales y en otros instrumentos). Hay temáticas en las que a pesar de que se utilicen preguntas sutiles, el encuestado se puede sentir molesto. Tal es el caso del desempleo, la homosexualidad, el sida, la prostitución, la pornografía, los anticonceptivos y las adicciones.

e) Las preguntas deben referirse preferentemente a un solo aspecto o una relación lógica. Por ejemplo, la pregunta: ¿acostumbra usted ver televisión y escuchar radio diariamente?, expresa dos aspectos y llega a confundir. Es necesario dividirla en dos preguntas, una relacionada con la televisión y otra vinculada con la radio. Otro ejemplo: ¿sus padres eran saludables?, es una pregunta problemática, además del concepto "saludable" (confuso), es imposible de responder en el caso

www.elosopanda.com | jamespoetrodriguez.com

de que la madre nunca se hubiera enfermado de gravedad y nunca hubiera sido hospitalizada y, en cambio, el padre hubiera padecido severos problemas de salud.

f) Las preguntas no habrán de inducir las respuestas. Se tienen que evitar preguntas tendenciosas o que dan pie a elegir un tipo de respuesta (directivas). Por ejemplo, ¿considera a nuestro compañero Ricardo Hernández como el mejor candidato para dirigir nuestro sindicato?, es una pregunta tendenciosa, pues induce la respuesta. Lo mismo que la pregunta: ¿los trabajadores argentinos son muy productivos? Se insinúa la respuesta en la pregunta. Resultaría mucho más conveniente interrogar: ¿qué tan productivos considera usted, en general, a los trabajadores argentinos? (y mostrar alternativas).

Ejemplo

¿Qué tan productivos considera usted, en general, a los trabajadores argentinos?

Sumamente productivos	Más bien productivos	Más bien improductivos	Sumamente improductivos

Otro ejemplo inconveniente sería: ¿piensa usted votar por tal partido político en las próximas elecciones? El participante nunca debe sentirse presionado. Un factor importante a considerar es la deseabilidad social (Furr, 2010), a veces las personas utilizan respuestas culturalmente aceptables. Por ejemplo, la pregunta: ¿le gustaría casarse?, podría inducir y forzar a más de una persona a responder de acuerdo con las normas de su comunidad. Resulta mejor cuestionar: ¿qué opina del matrimonio?, y más adelante inquirir sobre sus anhelos y expectativas al respecto. Una interrogante como: ¿acostumbra leer el periódico?, puede llevarnos a respuestas socialmente válidas: "sí, lo leo a diario, yo leo mucho" (cuando no es cierto). Es mejor preguntar: ¿suele tener tiempo para leer el periódico?, ¿con qué frecuencia?

g) Las preguntas no pueden apoyarse en instituciones, ideas respaldadas socialmente ni en evidencia comprobada. Es también una manera de inducir respuestas. Por ejemplo, la pregunta: la Organización Mundial de la Salud ha realizado diversos estudios y concluyó que el tabaquismo provoca diversos daños al organismo, ¿considera usted que fumar es nocivo para su salud? Esquemas del tipo: "la mayoría de las personas opinan que…", "la Iglesia considera…", "los padres de familia piensan que…", etc., no deben anteceder a las preguntas, ya que influyen y sesgan las respuestas.

h) Es aconsejable evitar preguntas que nieguen el asunto que se interroga. Por ejemplo: ¿qué niveles de la estructura organizacional *no* apoyan el proceso de calidad? Es mejor preguntar sobre qué niveles sí apoyan el proceso. O bien: ¿qué *no* le agrada de este centro comercial?, es preferible cuestionar: ¿qué le desagrada de este centro comercial? Tampoco es conveniente incluir dobles negaciones (son positivas, pero suelen confundir): ¿considera que la mayoría de las mujeres casadas preferiría *no* trabajar si *no* tuviera presión económica? Mejor se redacta de manera positiva.

i) No deben hacerse preguntas racistas o sexistas ni que ofendan a los participantes. Es obvio, pero no está de más recalcarlo. Se recomienda también sortear las preguntas *con fuerte carga emocional o muy complejas*, que más bien son preguntas para entrevistas cualitativas (por ejemplo: ¿cómo era la relación con su ex marido? —aunque una escala completa puede ser la solución— o ¿qué siente usted sobre la muerte de su hijo?).

j) En las preguntas con varias categorías de respuesta, y cuando el entrevistado sólo tiene que elegir una, llega a ocurrir que el orden en que se presentan dichas opciones afecta las respuestas de los participantes (por ejemplo, que tiendan a favorecer a la primera o a la última opción de respuesta). Entonces resulta conveniente rotar el orden de lectura de las respuestas a elegir de manera proporcional. Por ejemplo, si preguntamos: ¿cuál de los siguientes cuatro candidatos presidenciales considera usted que logrará disminuir verdaderamente la inflación? Veinticinco por ciento de las veces (o una de cada cuatro ocasiones) que se haga la pregunta se menciona primero al candidato *A*, 25% se menciona primero al candidato *B*, *25%* al candidato *C* y el restante 25% al

candidato *D*. Asimismo, cuando las alternativas son demasiadas es más difícil responder, por ello es conveniente limitarlas a las mínimas necesarias.

En la tabla 9.4 se relacionan los problemas que surgen al generar preguntas, adaptada de Creswell (2005).

Tabla 9.4 Ejemplos de algunos problemas al elaborar preguntas

Problema	Ejemplo de pregunta problemática	Mejora a la pregunta
Pregunta confusa por la vaguedad de los términos	¿Votará en las próximas elecciones?	Precisar términos: En las próximas elecciones del 10 de noviembre para elegir alcalde de Monterrey, ¿piensa ir a votar?
Dos o más conceptos o dos preguntas en una sola	¿Qué tan satisfecho está usted con el servicio de comedor y el servicio médico que se ofrece en la empresa?	Una pregunta por concepto: ¿Qué tan satisfecho está usted con el servicio de comedor que se ofrece en la empresa? ¿Qué tan satisfecho está usted con el servicio médico que se ofrece en la empresa?
Demasiadas palabras	Como usted sabe, el próximo 10 de noviembre se celebrarán elecciones locales en este municipio de Cortazar para elegir alcalde, en esa fecha: ¿piensa usted acudir a las urnas a emitir su voto por el candidato que considera será el mejor alcalde para el municipio?	Reducir términos: En las próximas elecciones del 10 de noviembre para elegir alcalde de Cortazar, ¿piensa ir a votar?
Pregunta negativa	¿Los estudiantes no deben portar o llevar armas a la escuela?	Cambiarla a neutral: ¿Los estudiantes deben o no portar armas en la escuela?
Contiene "jerga lingüística"	¿Qué tan "chida" o "padre" es la relación con su empresa?	Eliminar dicha jerga: ¿Qué tan orgulloso se encuentra usted de trabajar en esta empresa?
Se traslapan las categorías de respuesta	¿Podría indicarme su edad? __18 – 19 __19 – 20 __20 – 21 __21 – 22	Lograr que las categorías sean mutuamente excluyentes: ¿Podría indicarme su edad? __18 – 19 __20 – 21 __22 – 23
Categorías de respuesta sin balance entre las favorables y las desfavorables (positivas y negativas)	¿En qué medida está usted satisfecho con su superior inmediato? ☐ Insatisfecho ☐ Medianamente satisfecho ☐ Satisfecho ☐ Sumamente satisfecho	Proporcionar equilibrio entre opciones favorables y desfavorables: ¿En qué medida está usted satisfecho con su superior inmediato? ☐ Sumamente insatisfecho ☐ Más bien insatisfecho ☐ Ni insatisfecho ni satisfecho ☐ Más bien satisfecho ☐ Sumamente satisfecho
Incongruencia entre la pregunta y las opciones de respuesta	¿En qué medida está usted satisfecho con su superior inmediato? ☐ Muy poco importante ☐ Poco importante	Generar categorías que coincidan con la pregunta: ¿En qué medida está usted satisfecho con su superior inmediato? ☐ Sumamente insatisfecho

(*continúa*)

Tabla 9.4 (*continuación*)

Problema	Ejemplo de pregunta problemática	Mejora a la pregunta
	☐ Medianamente importante ☐ Importante ☐ Muy importante	☐ Más bien insatisfecho ☐ Ni insatisfecho ni satisfecho ☐ Más bien satisfecho ☐ Sumamente satisfecho
Sólo una parte de los participantes pueden entender la pregunta	¿Cuál es el género y marca de bebida etílica que acostumbra adquirir con un mayor índice de frecuencia en sus compras?	Simplificar términos: ¿Cuál es el tipo de bebida alcohólica y de qué marca acostumbra comprar con mayor frecuencia?
Utilización de términos en otro idioma	¿Qué efectos tuvo en esta empresa el *downsizing*?	Traducir términos: ¿Qué efectos tuvo en esta empresa la reducción de empleados?
La pregunta puede ser inadecuada para parte de la población	¿Cómo le afectó el incremento en la tasa impositiva para empleados gubernamentales?	Agregar preguntas que segmenten a la población: ¿Actualmente trabaja? ☐ Sí ☐ No ¿Trabaja usted en... ☐ empresa? ☐ por cuenta propia (independiente)? ☐ gobierno? Entonces, a quienes pertenezcan a la última categoría, se les pregunta: ¿Cómo le afectó el incremento en la tasa impositiva para empleados gubernamentales?

En relación con cada pregunta del cuestionario, León y Montero (2003) sugieren que se reflexione sobre si es necesaria la pregunta. ¿Es lo suficientemente concreta? ¿Responderán los participantes sinceramente?

¿Cómo deben ser las primeras preguntas de un cuestionario?

En algunos casos es conveniente iniciar con preguntas neutrales o fáciles de contestar, para que el participante se adentre en la situación. No se recomienda comenzar con preguntas difíciles o muy directas. Imaginemos un cuestionario diseñado para obtener opiniones en torno al aborto que empiece con una pregunta poco sutil como: ¿está de acuerdo con que se legalice el aborto en este país? Sin lugar a dudas, sería un fracaso. Los primeros cuestionamientos deben resultar interesantes para los participantes. A veces incluso, pueden ser divertidos (por ejemplo, en la investigación de la moda y la mujer mexicana que se verá en la cuarta parte del libro sobre modelos mixtos, al comenzar a inquirir sobre los tipos de prendas que compraban las participantes, la primera pregunta fue: ¿sueles ponerte una piyama para dormir?, cuestionamiento que resultó sumamente divertido y provocó hilaridad, logrando relajar a las encuestadas. Desde luego, la pregunta la hicieron mujeres entrevistadoras jóvenes).

A veces los cuestionarios comienzan con las preguntas demográficas ya mencionadas, pero en otras ocasiones es mucho mejor hacer este tipo de preguntas al final del cuestionario, particularmente en casos donde los participantes puedan sentir que se comprometen si responden el cuestionario.

Cuando construimos un cuestionario, es indispensable que pensemos en cuáles son las preguntas ideales para iniciar. Éstas deberán lograr que el sujeto se concentre en el cuestionario. Al respecto, se

sugiere el procedimiento de "embudo" en la presentación de las preguntas: ir de las más generales a las más específicas. Una característica fundamental de un cuestionario es que las preguntas importantes nunca deben ir al final.

¿De qué está formado un cuestionario?

Además de las preguntas y categorías de respuestas, un cuestionario está formado básicamente por: portada, introducción, instrucciones insertas a lo largo del contenido y agradecimiento final.

Portada

Esta incluye la carátula; en general, debe ser atractiva gráficamente para favorecer las respuestas. Debe incluir el nombre del cuestionario y el logotipo de la institución que lo patrocina. En ocasiones se agrega un logotipo propio del cuestionario o un símbolo que lo identifique.

Introducción

Debe incluir:

- Propósito general del estudio.
- Motivaciones para el sujeto encuestado (importancia de su participación).
- Agradecimiento.
- Tiempo aproximado de respuesta (un promedio o rango).
- Espacio para que firme o indique su consentimiento (a veces se incluye al final o en ocasiones es innecesario).
- Identificación de quién o quiénes lo aplican.
- Explicar brevemente cómo se procesarán los cuestionarios y una cláusula de confidencialidad del manejo de la información individual.
- Instrucciones iniciales claras y sencillas (cómo responder en general, con ejemplos si se requiere).

Cuando el cuestionario se aplica mediante entrevista, la mayoría de tales elementos son explicados por el entrevistador. El cuestionario debe ser y parecer corto, fácil y atractivo.

A continuación, se presenta un ejemplo de carta introductoria y otro de instrucciones generales para responder al cuestionario.[16]

Ejemplo

Carta introductoria

Buenos días (tardes):
Estamos trabajando en un estudio que servirá para elaborar una tesis profesional acerca de la biblioteca de la Universidad de Celaya.

Quisiéramos pedir tu ayuda para que contestes algunas preguntas que no te tomarán mucho tiempo. Tus respuestas serán confidenciales y anónimas. No hay preguntas delicadas.

Las personas que fueron seleccionadas para el estudio no se eligieron por su nombre sino al azar. Las opiniones de todos los encuestados serán sumadas e incluidas en la tesis profesional, pero nunca se comunicarán datos individuales.

Te pedimos que contestes este cuestionario con la mayor sinceridad posible. No hay respuestas correctas ni incorrectas.

Lee las instrucciones cuidadosamente, ya que existen preguntas en las que sólo se puede responder a una opción; otras son de varias opciones y también se incluyen preguntas abiertas.

¡Muchas gracias por tu colaboración!

[16] El lenguaje es claramente el que se utiliza en México. Desde luego, cada quien debe adaptarlo a su propia cultura.

Ejemplo

Instrucciones de un cuestionario

ENCUESTA DEL CLIMA ORGANIZACIONAL

INSTRUCCIONES

Emplee un lápiz o un bolígrafo de tinta negra para responder el cuestionario. Al hacerlo, piense en lo que sucede la mayoría de las veces en su trabajo.

No hay respuestas correctas o incorrectas. Simplemente reflejan su opinión personal.

Todas las preguntas tienen cinco opciones de respuesta. Elija la que mejor describa lo que piensa usted. Solamente una opción.

Marque con claridad la opción elegida con una cruz o tache, o bien, una "paloma" (símbolo de verificación). Recuerde: NO se deben marcar dos opciones. Marque así:

☒ ☑

Si no puede contestar una pregunta o si la pregunta no tiene sentido para usted, por favor pregúntele a la persona que le entregó este cuestionario y le explicó la importancia de su participación.

CONFIDENCIALIDAD

Sus respuestas serán anónimas y absolutamente confidenciales. Los cuestionarios serán procesados por personas externas. Además, como usted puede ver, en ningún momento se le pide su nombre.

De antemano: ¡muchas gracias por su colaboración!

Algunas apelaciones que podemos utilizar en la introducción se muestran en la tabla 9.5.[17]

Tabla 9.5 Ejemplos de apelaciones para incentivar la participación

Apelación	Ejemplo
Incentivo	"Al responder, usted recibirá..." (dinero, un obsequio, un boleto, etcétera).
Altruismo	"Los resultados servirán para resolver...", "El estudio ayudará a..." (problema social, mejora en la calidad de vida, solventar una necesidad comunitaria, etcétera).
Autoconcepto de la persona	"Usted es una de las pocas personas que puede señalar ciertas cuestiones...". "Debido a su experiencia (pericia, importancia, conocimientos, etc.) usted puede... y por ello le solicitamos..." (su opinión calificada, etcétera).
Interés por el conocimiento	"Le enviaremos una copia de los resultados..."
Intereses profesionales	"Los resultados serán útiles para conocer temas importantes en nuestra profesión..."
Ayuda-auxilio	"Necesitamos su apoyo para conocer...", "Los jóvenes requieren de ayuda para..."
Autoridad	Introducción acompañada de la firma de un líder o persona reconocida. O bien: "Doña Pola Castelán nos ha pedido que hagamos esta encuesta para conocer el problema de los niños...", "El científico...", "El empresario..."
Agradecimiento	"La comunidad de... estará muy agradecida por..."

También se insertan instrucciones a lo largo del cuestionario (normalmente con otra fuente tipográfica o bien en cursivas, para distinguirlas de las preguntas y respuestas), las cuales nos indican cómo contestar. Por ejemplo:

–¿Tiene este ejido o esta comunidad ganado, aves o colmenas que sean de propiedad colectiva? (*Marque con una cruz su respuesta*).

☐ Sí (continúe) ☐ No (pase a la pregunta 30)

–¿Se ha obtenido la cooperación de todo el personal o de la mayoría de éste para el proyecto de mejora de la calidad?

1 Sí (*pase a la pregunta 26*) 2 No (*pase a la pregunta 27*)

[17] Basada en Mertens (2010).

Recordemos que en ocasiones se presentan tarjetas con las opciones de respuestas y se instruye al entrevistador para que las muestre a los participantes. Por ejemplo:

–Hablando de la mayoría de sus proveedores, en qué medida conoce usted… (*mostrar la tarjeta uno y marcar la respuesta en cada caso*).

Ejemplo

De pregunta con tarjeta de respuestas

	Completamente (5)	Bastante (4)	Regular (3)	Poco (2)	Nada (1)
• ¿Sus políticas de venta?					
• ¿Sus finanzas (estados financieros)?					
• ¿Los objetivos de su área de ventas?					
• ¿Sus programas de capacitación para vendedores?					
• ¿Número de empleados de su área de ventas?					
• ¿Problemas laborales?					
• ¿Sus métodos de producción?					
• ¿Otros clientes de ellos?					
• ¿Su índice de rotación de personal?					

Las instrucciones son tan importantes como las preguntas y es necesario que sean claras para los usuarios a quienes van dirigidas.

Agradecimiento final

Aunque haya agradecido de antemano, vuelva a agradecer la participación.

Formato, distribución de instrucciones, preguntas y categorías

Las preguntas deben estar organizadas para que sea más fácil de responder el cuestionario. Es importante asegurarnos de numerar páginas y preguntas.

La manera en que pueden distribuirse preguntas, categorías de respuesta e instrucciones es variada. Algunos prefieren colocar las preguntas a la izquierda y las respuestas a la derecha, con lo que se tendría un formato como el siguiente:

Ejemplo

Modelo de formato de distribución de preguntas

¿Considera a su jefe o superior inmediato como su amigo?	☐ Definitivamente sí	☐ Sí	☐ No	☐ Definitivamente no
¿Cuando tiene problemas se siente apoyado por su jefe o superior inmediato?	☐ Definitivamente sí	☐ Sí	☐ No	☐ Definitivamente no
¿Considera que su jefe o superior inmediato le orienta adecuadamente en su trabajo?	☐ Definitivamente sí	☐ Sí	☐ No	☐ Definitivamente no
¿Tiene una buena impresión de su jefe o superior inmediato?	☐ Definitivamente sí	☐ Sí	☐ No	☐ Definitivamente no

Otros dividen el cuestionario por secciones de preguntas y utilizan un formato horizontal.

Ejemplo

Modelo de formato horizontal

Presentación

PREGUNTAS SOBRE EL SUPERIOR INMEDIATO

¿Considera a su jefe o superior inmediato como su amigo?

☐ Definitivamente sí ☐ Sí ☐ No ☐ Definitivamente no

¿Cuando tiene problemas se siente apoyado por su jefe o superior inmediato?

☐ Definitivamente sí ☐ Sí ☐ No ☐ Definitivamente no

¿Considera que su jefe o superior inmediato le orienta adecuadamente en su trabajo?

☐ Definitivamente sí ☐ Sí ☐ No ☐ Definitivamente no

¿Tiene una buena impresión de su jefe o superior inmediato?

☐ Definitivamente sí ☐ Sí ☐ No ☐ Definitivamente no

PREGUNTAS SOBRE MOTIVACIÓN

Otros combinan diversas posibilidades, distribuyendo preguntas que miden la misma variable a través de todo el cuestionario. Cada quien es capaz de utilizar el formato que desee o juzgue conveniente, lo importante es que en su totalidad sea comprensible para el usuario: que las instrucciones, preguntas y respuestas se diferencien; que el formato no resulte visualmente tedioso y se lea sin dificultad.

Hoy en día es común elaborar cuestionarios electrónicos para dispositivos portátiles de diferentes tipos como laptops, tabletas, teléfonos móviles, etc., así como para páginas web y blogs que contienen fotografías, dibujos, secuencias de video y música. Son sumamente atractivos y existen decenas de programas para diseñarlos.

¿De qué tamaño debe ser un cuestionario?

No existe una regla al respecto, pero si el cuestionario es muy corto se pierde información y si resulta largo llega a ser tedioso. En este último caso, las personas se negarían a responder o, al menos, lo contestarían en forma incompleta. La abuela doña Margarita Castelán Sampieri repetía el refrán: "lo bueno y breve, doblemente bueno". El tamaño depende del número de variables y dimensiones a medir, el interés de los participantes y la manera como se administre (de este punto se hablará en el siguiente apartado). Cuestionarios que toman más de 35 minutos suelen resultar fatigosos, a menos que los sujetos estén muy motivados para contestar (por ejemplo, cuestionarios de personalidad o cuestionarios para obtener un trabajo). Una recomendación que ayuda a evitar un cuestionario más largo de lo requerido es: *no hacer preguntas innecesarias o injustificadas.*

¿Cómo se codifican las preguntas abiertas?

Las preguntas abiertas se codifican una vez que conocemos todas las respuestas de los participantes a los cuales se les aplicaron, o al menos las principales tendencias de respuestas en una muestra de los cuestionarios aplicados. Es importante anotar que esta actividad es similar a "cerrar" una pregunta abierta por medio de la prueba piloto, pero el producto es diferente. En este caso, con la codificación de preguntas abiertas se obtienen ciertas categorías que representan los resultados finales.

El procedimiento consiste en encontrar y dar nombre a los patrones generales de respuesta (respuestas similares o comunes), listar estos patrones y después asignar un valor numérico o un símbolo a cada patrón. Así, un patrón constituirá una categoría de respuesta. Para cerrar las preguntas abiertas se sugiere el siguiente procedimiento:

1. Seleccionar determinado número de cuestionarios mediante un método adecuado de muestreo, que asegure la representatividad de los participantes investigados.
2. Observar la frecuencia con que aparece cada respuesta en las preguntas.
3. Elegir las respuestas que se presentan con mayor frecuencia (patrones generales de respuesta).
4. Clasificar las respuestas elegidas en temas, aspectos o rubros, de acuerdo con un criterio lógico, cuidando que sean mutuamente excluyentes.
5. Darle un nombre o título a cada tema, aspecto o rubro (patrón general de respuesta).
6. Asignarle el código a cada patrón general de respuesta.

Por ejemplo, en una investigación sobre mejora de la calidad se hizo una pregunta abierta: ¿de qué manera la alta gerencia busca obtener la cooperación del personal para el desarrollo del proyecto de calidad?

Las respuestas fueron múltiples, pero se encontraron los patrones generales de respuesta que se muestran en el ejemplo.

Ejemplo

De codificación de preguntas abiertas

Códigos	Categorías (patrones o respuestas con mayor frecuencia de mención)	Frecuencia de mención
1	Involucrar al personal y comunicarse con él	28
2	Motivación e integración	20
3	Capacitación en general	12
4	Incentivos/recompensas	11
5	Difundir el valor "calidad" o la filosofía de la empresa	7
6	Grupos o sesiones de trabajo	5
7	Posicionamiento del área de calidad o equivalente	3
8	Sensibilización en grupo	2
9	Desarrollo de la calidad de vida en el trabajo	2
10	Incluir aspectos de calidad en el manual de inducción	2
11	Poner énfasis en el cuidado de la maquinaria	2
12	Trabajar bajo un buen clima laboral	2
13	Capacitación "en cascada"	2
14	Otras	24

Como varias categorías o diversos patrones tenían únicamente dos frecuencias, se redujeron a sólo seis, tal como se ejemplifica a continuación.

Ejemplo

De reducción o agrupamiento de categorías

Códigos	Categorías (frecuencias)
1	Involucrar al personal y comunicarse con él (28)
2	Motivación e integración/mejoramiento del clima laboral (22)
3	Capacitación (14)
4	Incentivos/recompensas (11)
5	Difundir el valor "calidad" o la filosofía de la empresa (7)
6	Grupos o sesiones de trabajo (7)
7	Otras (33)

Al "cerrar" preguntas abiertas y codificarlas, debe tenerse en cuenta que un mismo patrón de respuesta puede expresarse con diferentes palabras. Por ejemplo, ante la pregunta: ¿qué sugerencias haría para mejorar al programa *Estelar*? Las respuestas: mejorar las canciones y la música, cambiar las canciones, incluir nuevas y mejores canciones, etc., se agruparían en la categoría o el patrón de respuesta *modificar la musicalización del programa.*

¿En qué contextos puede administrarse o aplicarse un cuestionario?

Los cuestionarios se aplican de dos maneras fundamentales: autoadministrado y por entrevista (personal o telefónica). En este apartado se comentan ambas y en el capítulo siete adicional "Recolección de los datos cuantitativos: segunda parte" (en el centro de recursos en línea: Material complementario → Capítulos), en la sección "Contextos de administración de los cuestionarios" se amplían algunos conceptos.

Centro de recursos en línea

1. Autoadministrado

Autoadministrado significa que el cuestionario se proporciona directamente a los participantes, quienes lo contestan. No hay intermediarios y las respuestas las marcan ellos. Pero la forma de autoadministración puede tener distintos contextos: individual, grupal o por envío (correo tradicional, correo electrónico y página web o equivalente).

En el caso individual, el cuestionario se entrega al participante y éste lo responde, ya sea que acuda a un lugar para hacerlo (como ocurre cuando se llena un formulario para solicitar empleo) o lo conteste en su lugar de trabajo, hogar o estudio. Por ejemplo, si los participantes fueran una muestra de directivos de laboratorios farmacéuticos de Bogotá, se acudiría a sus oficinas y se les entregarían los cuestionarios. Los ejecutivos se administrarían ellos mismos el cuestionario y esperaríamos a que lo respondieran o lo recogeríamos otro día. El reto de esta última situación es lograr que los participantes devuelvan el cuestionario contestado completamente. Resulta conveniente que quien lo entregue posea habilidades para relacionarse con las personas, sea asertivo, y además se caracterice por una elevada persistencia. En nuestra experiencia, en distintos países de Iberoamérica, jóvenes de ambos géneros con buena capacidad comunicativa logran porcentajes de recuperación por encima de 90% en tiempos aceptables (una semana o menos). Y no es necesario que sean físicamente atractivos (aunque ayuda), más bien el éxito reside en su motivación y tenacidad. Asimismo, el mayor costo o gasto de esta clase de administración de los cuestionarios lo representa su distribución y recolección.

En el segundo caso, se reúne a los participantes en grupos (a veces pequeños —cuatro a seis personas—, otras en grupos intermedios —entre siete y 20 sujetos—, incluso en grupos grandes de 21 a 40 individuos). Por ejemplo, empleados (en encuestas de clima organizacional es muy común juntar a grupos de 25, entregarles el cuestionario, introducir al propósito del estudio y al instrumento, responder dudas y pedirles que al concluir lo depositen en una urna sellada, para mantener la confidencialidad), padres de familia (en reuniones escolares), televidentes (cuando asisten a un foro televisivo), alumnos (en sus salones de clase), etc. Es tal vez la forma más económica de aplicar un cuestionario.

Cuando se envía a los participantes para autoadministración, puede usarse correo postal público o mensajería privada, o bien, correo electrónico. También se les puede pedir que ingresen a una página web o equivalente para responderlo.

Los cuestionarios para autoadministración deben ser particularmente atractivos (a colores, en papel especial, con diseño original, etc.; si el presupuesto lo permite).

2. Por entrevista personal

Las *entrevistas* implican que una persona calificada (entrevistador) aplica el cuestionario a los participantes; el primero hace las preguntas a cada entrevistado y anota las respuestas. Su papel es crucial, resulta una especie de filtro.

El primer contexto que se revisará de una entrevista es el personal ("cara a cara").

Normalmente se tienen varios entrevistadores, quienes deberán estar capacitados en el arte de entrevistar y conocer a fondo el cuestionario. No deberán sesgar o influir en las respuestas, por ejem-

plo, reservarse de expresar aprobación o desaprobación respecto de las respuestas del entrevistado, reaccionar de manera ecuánime cuando los participantes se alteren, contestar con gestos ambiguos cuando los sujetos busquen generar una reacción en ellos, etc. Su propósito es llevar a buen fin cada entrevista, evitando que decaiga la concentración e interés del participante, además de orientarlo en el tránsito del instrumento. Las explicaciones que proporcione deberán ser breves pero suficientes. Tiene que ser neutral, pero cordial y servicial. Asimismo, es muy importante que transmita a todos los participantes que no hay respuestas correctas o equivocadas. Por otra parte, su proceder debe ser lo más homogéneo posible (mismos señalamientos, presentación uniforme, etc.). Con respecto a las instrucciones del cuestionario, algunas son para el entrevistado y otras para el entrevistador. Este último debe recordar que al inicio se comenta: el propósito general del estudio, las motivaciones y el tiempo aproximado de respuesta, agradeciendo de antemano la colaboración.

El cuestionario por entrevista es el que consigue un mayor porcentaje de respuestas a las preguntas, su estimación es de 80 a 85% (León y Montero, 2003).

En relación con el perfil de los entrevistadores no hay un consenso; por ejemplo, Corbetta (2003) sugiere que sean mujeres casadas, amas de casa, de mediana edad, diplomadas y de clase media. León y Montero (2003) recomiendan que sean siempre profesionales. En nuestra experiencia, el tipo de entrevistador depende del tipo de persona entrevistada. Por ejemplo, que pertenezca a un nivel socioeconómico similar a la mayoría de la muestra, sea joven y haya cursado asignaturas de investigación, que posea facilidad de palabra y capacidad de socializar. Se reitera que los estudiantes de ambos géneros funcionan muy bien. Por tanto, es claro que para este fin deben evitarse personas inseguras o excesivamente tímidas.

La entrevista cuantitativa es diferente de la cualitativa.[18] En relación a la primera, Moule y Goodman (2009) y Rogers y Bouey (2005), mencionan las siguientes características:

- ***a***) El principio y final de la entrevista se definen con claridad. De hecho, tal definición se integra en el cuestionario.
- ***b***) El mismo instrumento se aplica a todos los participantes, en condiciones lo más similares posible (estandarización).
- ***c***) Se busca que sea individual, sin la intrusión de otras personas que pueden opinar o alterar de alguna manera la entrevista.
- ***d***) Es poco a nada anecdótica (aunque en algunos casos es recomendable que el entrevistador anote cuestiones fuera de lo común como ciertas reacciones y negativas a responder).
- ***e***) La mayoría de las preguntas suelen ser cerradas, con mínimos elementos rebatibles, ampliaciones y sondeos.
- ***f***) El entrevistador y el propio cuestionario controlan el ritmo y la dirección de la entrevista.
- ***g***) El contexto social no es un elemento a considerar, lo es solamente el ambiental.
- ***h***) El entrevistador procura que su patrón de comunicación sea similar (su lenguaje, instrucciones, etcétera).

Asimismo, la capacitación de entrevistadores debe incluir cuestiones de comunicación no verbal básicas (control de gestos, manejo de silencios, etc.), además de todos los puntos que se revisaron anteriormente.

Cabe señalar que, cuando se trata de *entrevista personal*, las condiciones del lugar donde se realice son importantes (oficina, hogar o casa-habitación, sitio público, como centro comercial, parque, escuela, etc.). Pero, en cualquier caso, se aconseja que se busque un lugar lo más discreto, silencioso y privado que sea posible.

En estas entrevistas es común mostrar visualmente las opciones de respuesta a los entrevistados, mediante tarjetas, en especial cuando se incluyen más de cinco o son complejas. Pongamos de ejemplo la siguiente tarjeta.[19]

[18] Las entrevistas cualitativas se revisan en el capítulo 14 "Recolección y análisis de los datos cualitativos".
[19] El ejemplo se ha simplificado por cuestiones de espacio, las opciones fueron obtenidas después de una prueba piloto.

www.elosopanda.com | jamespoetrodriguez.com

Ejemplo

Tarjeta para mostrar al entrevistado cuando hay diversas opciones de respuesta

¿Cuáles considera usted que son los tres principales problemas en este municipio?

- ☐ Pandillerismo
- ☐ Venta de drogas, narcomenudeo
- ☐ Pobreza
- ☐ Corrupción de funcionarios de la alcaldía
- ☐ Desempleo
- ☐ Falta de vivienda
- ☐ Falta de infraestructura (calles, puentes, etcétera)
- ☐ Empleo mal remunerado
- ☐ Inseguridad en las calles
- ☐ Problemas en la recolección de basura
- ☐ Escasez de agua
- ☐ Carencia de servicios de salud

Hoy en día hay varios sistemas para sustituir al cuestionario de lápiz y papel, como la CAPI (*Computer-Assisted Personal Interviewing*), en la que el entrevistador muestra al participante una computadora portátil (*notebook* o *tableta*) o un teléfono móvil que contiene el cuestionario y este último responde guiado por el primero o directamente aquel anota las respuestas. Se usa teclado, sistema al tacto o pluma electrónica.

Casi siempre las **entrevistas** son individuales, aunque podrían aplicarse a un grupo pequeño (si ésta fuera la unidad de análisis). Es decir, el cuestionario lo responden entre todos sus miembros o parte de ellos (por ejemplo, cuestionarios para parejas o una familia, o un área de una empresa).

En entrevistas Si la pregunta se va a presentar oralmente, no debe contener más de cinco opciones de respuesta, ya que por encima de este límite se suelen olvidar las primeras.

Los cuestionarios administrados usando dispositivos tienen la ventaja de que los datos se capturan y agregan a la base de datos de forma automática, de manera que en cualquier momento podemos hacer un corte y efectuar toda clase de análisis (ver tendencias, evaluar funcionamiento del instrumento, etc.). La desventaja es obvia: el costo, el cual es muy difícil de absorber por parte de un estudiante o maestro e incluso una institución.

Idealmente, después de una entrevista se puede preparar un informe que indique: si el participante se mostraba sincero, la manera como respondió, el tiempo que duró la entrevista, el lugar donde se realizó, las características del entrevistado, los contratiempos que se presentaron y la forma en que se desarrolló la entrevista, así como otros aspectos relevantes.

3. Por entrevista telefónica

Obviamente, la diferencia con el anterior tipo de entrevista es el medio de comunicación, que en este caso es el teléfono (hogar, oficina, móvil o celular). Las entrevistas telefónicas son la forma más rápida y económica de realizar una encuesta.

Las habilidades requeridas de parte de los entrevistadores son parecidas a las de la entrevista personal, excepto que éstos no tienen que confrontarse "cara a cara" con los participantes (no importa la vestimenta ni el aspecto físico, pero sí la voz; su modulación y claridad son fundamentales). El nivel de rechazo varía de acuerdo con el entorno social y tiempo en que se realice ("saturación telefónica" de campañas comerciales y políticas de telemercadeo, inseguridad pública, etcétera).

Una ventaja de este método reside en que se puede acceder a barrios inseguros y a hogares exclusivos donde se limita el ingreso, así como a lugares geográficamente lejanos.

La desventaja de las encuestas telefónicas es que están limitadas a unas cuantas preguntas o no se pueden efectuar mediciones complejas de variables o profundizar en ciertos temas. Pero una vez más, los datos se capturan y agregan a la base de datos de forma automática y se pueden hacer cortes de la información de manera inmediata y realizar toda clase de análisis.

Algunas recomendaciones para las entrevistas telefónicas y el tema de sistemas automatizados para realizarlas como el CATI (Computer-Assisted Telephone Interviewing), se revisan en el capítulo 7 adicional, "Recolección de los datos cuantitativos: segunda parte" (que se puede descargar del centro de recursos en línea).

Centro de recursos en línea

Cuando se realizan entrevistas personales en el hogar o telefónicas se debe tomar en cuenta el *horario.* Ya que si efectuamos la visita o hablamos por teléfono sólo a una hora (digamos en la mañana), nos encontraremos con unos cuantos subgrupos de la población (por ejemplo, amas de casa).

Una variación de la administración de cuestionarios por teléfono es la siguiente: en un programa radial, televisivo o vía internet se solicita la opinión o respuesta de los receptores a una pregunta o algunas cuantas preguntas, éstos deben marcar un número telefónico y contestar las opciones de respuesta con las que concuerden más. El problema de estas encuestas reside en la muestra, que desde luego no es probabilística, sino que se conforma de voluntarios que cubren dos condiciones: tener teléfono y estar viendo o escuchando la emisión del programa. Este proceder nos conduce más que a un estudio, a un sondeo, lo cual no es un error en sí, lo grave es que se pretenda generalizar los resultados a una población.

Algunas consideraciones adicionales para la administración del cuestionario

Cuando se tiene población analfabeta, con niveles educativos bajos o niños que apenas comienzan a leer o no dominan la lectura, el método más conveniente de administración de un cuestionario es por entrevista. Aunque existen algunos cuestionarios muy gráficos que usan escalas sencillas para las opciones de respuestas. Como en el siguiente ejemplo.

Ejemplo

Con trabajadores de niveles de lectura básica se recomienda utilizar entrevistas o cuestionarios autoadministrados sencillos que se apliquen en grupos, proporcionándoles asesoría.

En algunos casos, con ejecutivos que difícilmente vayan a dedicarle a un solo asunto más de 20 minutos, se pueden utilizar cuestionarios autoadministrados o entrevistas telefónicas. Con estudiantes suelen funcionar los cuestionarios autoadministrados.

Algunas asociaciones realizan encuestas por correo y ciertas empresas envían cuestionarios a sus empleados mediante el servicio interno de mensajería o por correo electrónico. Cuando el cuestionario contiene unas cuantas preguntas (su administración toma entre cuatro y cinco minutos), la entrevista telefónica es una buena alternativa.

Ahora bien, sea cual fuere la forma de administración, siempre debe haber uno o varios supervisores que verifiquen que los cuestionarios se están aplicando correctamente.

La elección del contexto para administrar el cuestionario deberá ser muy cuidadosa y dependerá del presupuesto disponible, el tiempo de entrega de los resultados, el planteamiento del problema, la naturaleza de los datos y el tipo de participantes (edad, nivel educativo, etcétera).

A continuación, en la tabla 9.6, se comparan esquemáticamente las principales formas de administración de cuestionarios.

www.elosopanda.com | jamespoetrodriguez.com

Tabla 9.6 Comparación de las principales formas de administración de cuestionarios

Método de administración	Tasa de respuesta	Presupuesto o costo (fuente que origina el mayor gasto)	Rapidez con que se administra	Profundidad de los datos obtenidos	Tamaño del cuestionario
Autoadministrado (individual)	Media	Medio (pago de recolectores)	Media	Alta	Cualquier tamaño razonable
Autoadministrado (grupal)	Alta	Bajo (sesiones)	Rápida	Alta	Cualquier tamaño razonable
Autoadministrado (envío por correo o paquetería)	Baja	Bajo por correo postal (envíos) Medio por paquetería (envíos)	Lenta	Alta	Cualquier tamaño razonable
Autoadministrado por correo electrónico o página web	Baja	Bajo (diseño electrónico)	Media	Alta	Cualquier tamaño razonable
Entrevista personal	Alta	Elevado (pago a entrevistadores y gastos de viaje)	Media	Alta	Cualquier tamaño razonable
Entrevista telefónica	Alta	Bajo (llamadas telefónicas y entrevistadores)	Rápida	Baja	Corto

Cuando los cuestionarios son muy complejos de contestar o de aplicar, suele utilizarse un manual que explica las instrucciones y cómo debe responderse o administrarse.

Escalas para medir las actitudes

4

Actitud Predisposición aprendida para responder coherentemente de manera favorable o desfavorable ante un objeto, ser vivo, actividad, concepto, persona o símbolo.

Una **actitud** es una predisposición aprendida para responder coherentemente de una manera favorable o desfavorable ante un objeto, ser vivo, actividad, concepto, persona o símbolo (Kassin, Fein y Markus, 2013; Devine y Plant, 2013; Oskamp y Schultz, 2009; y Fishbein y Ajzen, 1975). Así, los seres humanos tenemos actitudes hacia muy diversos objetos, símbolos, etc.; por ejemplo, actitudes hacia el aborto, la política económica, la familia, un profesor, diferentes grupos étnicos, la ley, nuestro trabajo, una nación específica, los osos, el nacionalismo, un proceso productivo, nosotros mismos, etcétera.

Las actitudes están relacionadas con el comportamiento que mantenemos en torno a los objetos o conceptos a que hacen referencia. Si mi actitud hacia el aborto es desfavorable, probablemente no abortaría o no participaría en un aborto. Si mi actitud es favorable a un partido político, lo más probable es que vote por él en las próximas elecciones. Desde luego, las actitudes sólo son un indicador de la conducta, pero no la conducta en sí. Por ello, las mediciones de actitudes deben interpretarse como "síntomas" y no como "hechos" (Padua, 2000). Si detecto que la actitud de un grupo hacia la contaminación es desfavorable, esto no significa que las personas estén tomando acciones para evitar contaminar el ambiente, aunque sí es un indicador de que pueden adoptarlas en forma paulatina. La actitud es como una "semilla" que en ciertas condiciones suele "germinar en comportamiento".

Las actitudes tienen diversas propiedades, entre las que destacan: dirección (positiva o negativa) e intensidad (alta o baja); estas propiedades forman parte de la medición.

5.4. Análisis y representación de datos

5.4.1. Organización e interpretación de datos cuantitativos y cualitativos – (Organización, procesamiento e interpretación de datos cuantitativos y cualitativos)

- Autor principal: **Roberto Hernández Sampieri**
- Libro recomendado: *Metodología de la Investigación*
- Autor de apoyo: **John W. Creswell**
- Obra: *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*
- **Autor complementario: Mario F. Triola**
- **Obra:** Estadística

5.4.2. Representación de datos mediante tablas y gráficos

- **Autor principal:** Roberto Hernández Sampieri
- **Libro recomendado:** *Metodología de la investigación*
- **Autor de apoyo:** Mario F. Triola
- **Obra:** *Estadística*

5.4.3. Técnicas estadísticas para el análisis de datos

- **Autor principal:** Roberto Hernández Sampieri
- **Libro recomendado:** Metodología de la Investigación
- **Autor de apoyo:** Mario F. Triola
- **Obra:** Estadística
- **Autor complementario:** Richard I. Levin & David S. Rubin
- **Obra:** Estadística para Administración y Economía

5.4.1. Organización e interpretación de datos cuantitativos y cualitativos

Para el desarrollo del subtema **Organización e interpretación de datos cuantitativos y cualitativos**, se consideran aquellos contenidos relacionados con la **preparación, organización, análisis e interpretación de la información obtenida durante una investigación**, diferenciando los procedimientos correspondientes a los datos cuantitativos y cualitativos. En el caso cuantitativo, se consideran el análisis de datos, el análisis preliminar, la presentación de estadísticas descriptivas y la interpretación de los resultados. En el caso cualitativo, se consideran la organización y preparación de los datos, su lectura general, codificación, identificación de temas, construcción de una interpretación y representación e interpretación de los datos. Estos contenidos permiten comprender el proceso mediante el cual los datos obtenidos se transforman en información organizada y posteriormente en resultados interpretables.

N.°	Libros	Capítulo	Contenido
1	*Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* — John W. Creswell y J. David Creswell	**Capítulo 8. Quantitative Methods**	**Data Analysis.** Análisis de los datos dentro de la investigación cuantitativa.
2	*Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* — John W. Creswell y J. David Creswell	**Capítulo 8. Quantitative Methods**	**Reporting the Descriptive Statistics.** Presentación de las estadísticas descriptivas obtenidas a partir de los datos cuantitativos.
3	*Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* — John W. Creswell y J. David Creswell	**Capítulo 8. Quantitative Methods**	**Preliminary Analysis.** Análisis preliminar de los datos antes de procedimientos estadísticos posteriores.
4	*Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* — John W. Creswell y J. David Creswell	**Capítulo 8. Quantitative Methods**	**Interpreting Results and Writing a Discussion Section.** Interpretación de los resultados obtenidos mediante el análisis cuantitativo.
5	*Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* — John W. Creswell y J. David Creswell	**Capítulo 9. Qualitative Methods**	**Data Analysis Procedures.** Procedimientos para el análisis de los datos cualitativos.
6	*Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* — John W. Creswell y J. David Creswell	**Capítulo 9. Qualitative Methods**	**Step 1. Organizing and Preparing the Data for Analysis.** Organización y preparación de los datos cualitativos para su análisis.

7	*Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* — John W. Creswell y J. David Creswell	**Capítulo 9. Qualitative Methods**	**Step 2. Reading Through All the Data / Step 3. Coding the Data.** Lectura general de los datos y codificación de la información.
8	*Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* — John W. Creswell y J. David Creswell	**Capítulo 9. Qualitative Methods**	**Step 4. Identifying Themes.** Identificación de temas a partir de la información codificada.
9	*Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* — John W. Creswell y J. David Creswell	**Capítulo 9. Qualitative Methods**	**Step 5. Developing a Story Line Interpretation / Step 6. Further Analyzing the Data Using an Analytic Framework.** Desarrollo de una interpretación y análisis posterior mediante un marco analítico.
10	*Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* — John W. Creswell y J. David Creswell	**Capítulo 9. Qualitative Methods**	**Step 7. Representing and Interpreting the Data.** Representación e interpretación de los datos cualitativos.

5.4.1. Organización e interpretación de datos cuantitativos y cualitativos

Para el desarrollo del subtema **5.4.1. Organización e interpretación de datos cuantitativos y cualitativos**, se consideran aquellos contenidos de Roberto Hernández Sampieri relacionados directamente con el **análisis, exploración, organización e interpretación de los datos cuantitativos y cualitativos**, evitando incorporar contenidos que corresponden específicamente a la **representación mediante tablas y gráficos (5.4.2)** o a las **técnicas estadísticas para el análisis de datos (5.4.3)**.

N.°	Libros	Capítulo	Página	Contenido
1	*Metodología de la Investigación* — Roberto Hernández Sampieri	Capítulo 10. **Análisis de datos cuantitativos**	p. 272	**¿Qué procedimiento se sigue para analizar cuantitativamente los datos?**
2	*Metodología de la Investigación* — Roberto Hernández Sampieri	Capítulo 10. **Análisis de datos cuantitativos**	p. 277	**Paso 3: explorar los datos**
3	*Metodología de la Investigación* — Roberto Hernández Sampieri	Capítulo 10. **Análisis de datos cuantitativos**	p. 282	**Estadística descriptiva para cada variable**
4	*Metodología de la Investigación* — Roberto Hernández Sampieri	Capítulo 10. **Análisis de datos cuantitativos**	p. 288	**¿Cómo se interpretan las medidas de tendencia central y de la variabilidad?**
5	*Metodología de la Investigación* — Roberto Hernández Sampieri	Capítulo 14. **Recolección y análisis de los datos cualitativos**	p. 418	**El análisis de los datos cualitativos**
6	*Metodología de la Investigación* — Roberto Hernández Sampieri	Capítulo 14. **Recolección y análisis de los datos cualitativos**	p. 420	**Reflexiones e impresiones durante la inmersión inicial**
7	*Metodología de la Investigación* — Roberto Hernández Sampieri	Capítulo 14. **Recolección y análisis de los datos cualitativos**	p. 421	**Reflexiones e impresiones durante la inmersión profunda**
8	*Metodología de la Investigación* — Roberto Hernández Sampieri	Capítulo 14. **Recolección y análisis de los datos cualitativos**	p. 422	**Análisis detallado de los datos**
9	*Metodología de la Investigación* — Roberto Hernández Sampieri	Capítulo 14. **Recolección y análisis de los datos cualitativos**	p. 425	**Bitácora de análisis**

5.4.1. Organización, procesamiento e interpretación de datos cuantitativos y cualitativos

Para el desarrollo del subtema **Organización, procesamiento e interpretación de datos cuantitativos y cualitativos**, se consideran aquellos contenidos relacionados con la identificación de los tipos de datos, su recopilación, organización, resumen, representación y descripción, así como con el uso de medidas estadísticas que permiten procesar e interpretar los datos cuantitativos obtenidos durante una investigación. En el caso de **Mario F. Triola**, se consideran principalmente los contenidos correspondientes a los capítulos 1, 2 y 3 de *Estadística*, debido a que abordan los tipos de datos, la recopilación de datos muestrales, las distribuciones de frecuencias, las representaciones gráficas y las medidas utilizadas para describir y explorar los datos.

N.°	Libros	Capítulo	Página	Contenido
1	*Estadística* – Mario F. Triola	**Capítulo 1. Introducción a la estadística**	**p. 1**	Introducción a la estadística y fundamentos para el trabajo con datos.
2	*Estadística* – Mario F. Triola	**1-2. Tipos de datos**	**p. 13**	**Identificación y clasificación de los tipos de datos**, aspecto fundamental para determinar cómo serán organizados y analizados.
3	*Estadística* – Mario F. Triola	**1-3. Recopilación de datos muestrales**	**p. 25**	Recopilación de datos muestrales como etapa previa a su organización, procesamiento y análisis.
4	*Estadística* – Mario F. Triola	**Capítulo 2. Exploración de datos con tablas y gráficas**	**p. 40**	Exploración, organización y presentación de datos mediante tablas y representaciones gráficas.
5	*Estadística* – Mario F. Triola	**2-1. Distribuciones de frecuencias para organizar y resumir datos**	**p. 42**	**Organización y resumen de datos mediante distribuciones de frecuencias**, frecuencias y porcentajes.
6	*Estadística* – Mario F. Triola	**2-2. Histogramas**	**p. 51**	Representación gráfica y exploración de la distribución de datos cuantitativos mediante histogramas.
7	*Estadística* – Mario F. Triola	**2-3. Gráficas que informan y gráficas que engañan**	**p. 57**	Interpretación adecuada de representaciones gráficas y reconocimiento de posibles representaciones engañosas.
8	*Estadística* – Mario F. Triola	**Capítulo 3. Descripción, exploración y comparación de datos**	**p. 80**	Descripción, exploración y comparación de conjuntos de datos.
9	*Estadística* – Mario F. Triola	**3-1. Medidas de tendencia central**	**p. 82**	Uso de medidas de tendencia central para **resumir y describir datos cuantitativos**.
10	*Estadística* – Mario F. Triola	**3-2. Medidas de variación**	**p. 97**	Medición de la variabilidad o dispersión de los datos para complementar su interpretación.

Autor: Gabriel Oxsa Quispe

capítulo

10 Análisis de datos cuantitativos

Al analizar los datos cuantitativos debemos recordar dos cuestiones: primero, que los modelos estadísticos son representaciones de la realidad, no la realidad misma; y segundo, los resultados numéricos siempre se interpretan en contexto, por ejemplo, un mismo valor de presión arterial no es igual en un bebé que en una persona de la tercera edad.

Roberto Hernández-Sampieri

Proceso de investigación cuantitativa

Paso 9 Analizar los datos

- Decidir el programa de análisis de datos que se utilizará.
- Explorar los datos obtenidos en la recolección.
- Analizar descriptivamente los datos por variable.
- Visualizar los datos por variable.
- Evaluar la confiabilidad, validez y objetividad de los instrumentos de medición utilizados.
- Analizar e interpretar mediante pruebas estadísticas las hipótesis planteadas (análisis estadístico inferencial).
- Realizar análisis adicionales.
- Preparar los resultados para presentarlos.

Objetivos de aprendizaje

Al terminar este capítulo, el alumno será capaz de:

1. Revisar el proceso para analizar los datos cuantitativos.
2. Reforzar los conocimientos estadísticos fundamentales.
3. Comprender las principales pruebas o métodos estadísticos desarrollados, así como sus aplicaciones y la forma de interpretar sus resultados.
4. Diferenciar la estadística descriptiva y la inferencial, la paramétrica y la no paramétrica.
5. Analizar la interrelación entre distintas pruebas estadísticas.

Síntesis

En el capítulo se presentan brevemente los principales programas computacionales de análisis estadístico que emplea la mayoría de los investigadores, así como el proceso fundamental para efectuar análisis cuantitativo. Asimismo, se comentan, analizan y ejemplifican las pruebas estadísticas más utilizadas. Se muestra la secuencia de análisis más común, con estadísticas descriptivas, análisis paramétricos, no paramétricos y multivariados. En la mayoría de estos análisis, el enfoque del capítulo se centra en los usos y la interpretación de los métodos, más que en los procedimientos de cálculo, debido a que los análisis se realizan con ayuda de una computadora.

Se realiza mediante programas computacionales como:

- SPSS®
- Minitab
- SAS
- STATS

Análisis de datos cuantitativos

Cuyo procedimiento es:

Fases

1. Seleccionar el programa estadístico para el análisis de datos
2. Ejecutar el programa
3. Explorar los datos: analizarlos y visualizarlos por variable del estudio
4. Se evalúa la confiabilidad y validez del o de los instrumentos escogidos
5. Se lleva a cabo análisis estadístico descriptivo de cada variable del estudio
6. Se realizan análisis estadísticos inferenciales respecto a las hipótesis planteadas
7. Se efectúan análisis adicionales
8. Se preparan los resultados para presentarlos

El análisis se realiza tomando en cuenta los niveles de medición de las variables y mediante la estadística, que puede ser

Descriptiva

- Distribución de frecuencias
- Medidas de tendencia central
  - Media
  - Mediana
  - Moda
- Medidas de variabilidad
  - Rango
  - Desviación estándar
  - Varianza
- Gráficas
- Puntuaciones *z* (en centro de recursos en línea)

Inferencia

- Sirve para estimar parámetros y probar hipótesis
- Se basa en la distribución muestral

- Análisis paramétrico
  - Coeficientes de correlación
  - Regresión lineal
  - Prueba *t*
  - Prueba de la diferencia de proporciones
  - Análisis de varianza
  - Análisis de covarianza (en centro de recursos en línea)
- Análisis no paramétrico
  - *Chi* cuadrada
  - Coeficientes de Spearman y Kendall
  - Coeficientes para tabulaciones cruzadas
- Análisis multivariados (centro de recursos en línea)

Nota: Este capítulo se complementa con uno adicional que se puede descargar del centro de recursos en línea, en: Material complementario → Capítulos → Capítulo 8, "Análisis estadístico: segunda parte"; junto con el documento 2, "Fórmulas y procedimientos estadísticos" y el apéndice 4 "Tablas anexas", que también pueden descargarse del mismo sitio .

Centro de recursos en línea

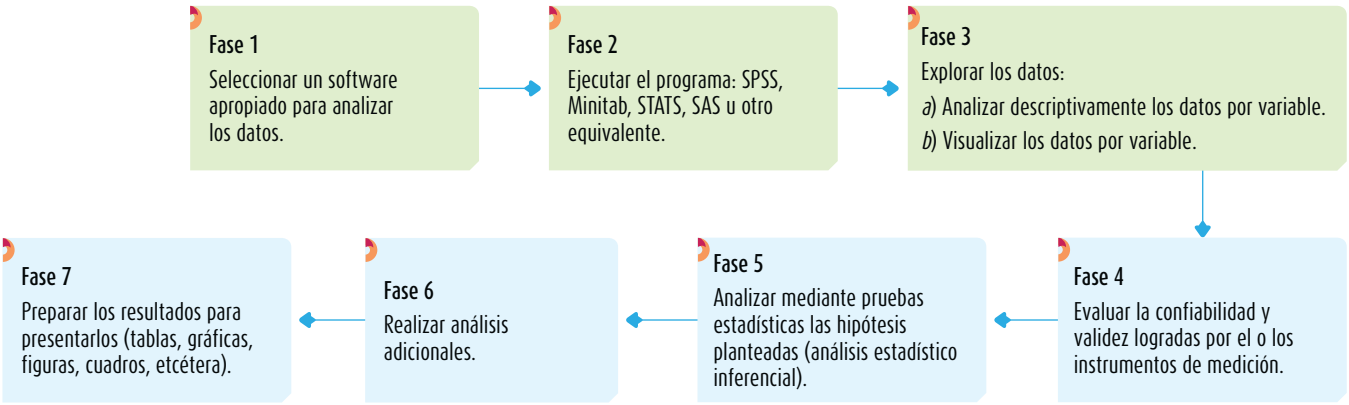
¿Qué procedimiento se sigue para analizar cuantitativamente los datos?

1 Una vez que los datos se han codificado, transferido a una matriz, guardado en un archivo y "limpiado" los errores, el investigador procede a analizarlos.

En la actualidad, el análisis cuantitativo de los datos se lleva a cabo *por computadora u ordenador*. Ya casi nadie lo hace de forma manual ni aplicando fórmulas, en especial si hay un volumen considerable de datos. Por otra parte, en la mayoría de las instituciones de educación media y superior, centros de investigación, empresas y sindicatos se dispone de sistemas de cómputo para archivar y analizar datos. De esta suposición parte el presente capítulo. Por ello, se centra en la *interpretación de los resultados de los métodos de análisis cuantitativo* y no en los procedimientos de cálculo.

El análisis de los datos se efectúa sobre la *matriz de datos* utilizando un *programa computacional*. El proceso de análisis se esquematiza en la figura 10.1. Posteriormente lo veremos paso a paso.

Figura 10.1 Proceso para efectuar análisis estadístico.



Paso 1: seleccionar un programa de análisis

Hay diversos programas para analizar datos. En esencia su funcionamiento es muy similar e incluyen las dos partes o segmentos que se mencionaron en el capítulo anterior: una parte de definiciones de las variables, que a su vez explican los datos (los elementos de la codificación ítem por ítem o indicador por indicador), y la otra parte, la matriz de datos. La primera parte es para que se comprenda la segunda. Las definiciones, desde luego, las prepara el investigador. Lo que éste hace, una vez recolectados los datos, es precisar los parámetros de la matriz de datos en el programa (nombre de cada variable en la matriz —que equivale a un ítem, reactivo, indicador, categoría o subcategoría de contenido u observación—, tipo de variable o ítem, ancho en dígitos, etc.) e introducir o capturar los datos en la matriz, la cual es como cualquier hoja de cálculo. Asimismo, recordemos que la matriz de datos tiene columnas (variables, ítems o indicadores), filas o renglones (casos) y celdas (intersecciones entre una columna y un renglón). Cada celda contiene un dato (que significa un valor de un caso en una variable). Supongamos que tenemos cuatro casos o personas y tres variables (género, color de cabello y edad); la matriz se vería como se muestra en la tabla 10.1.

Tabla 10.1 Ejemplo de matriz de datos con tres variables y cuatro casos

Caso	Columna 1 (género)	Columna 2 (color de pelo)	Columna 3 (edad)
1	1	1	35
2	1	1	29
3	2	1	28
4	2	4	33

La codificación (especificada en la parte de las definiciones de las variables o columnas que corresponden a ítems) sería:

- Género (1 = masculino y 2 = femenino).
- Color de cabello (1 = negro, 2 = castaño, 3 = pelirrojo, 4 = rubio).
- Edad (dato "bruto o crudo" en años).

De esta forma, si se lee por renglón o fila (caso), de izquierda a derecha, la primera celda indica un hombre (1); la segunda, de cabello negro (1); y la tercera, de 35 años (35). En el segundo renglón, un hombre de cabello negro y 29 años. La tercera fila, una mujer de cabello color negro, con 28 años. La cuarta fila (caso número cuatro) nos señala una mujer (2), rubia (4) y de 33 años (33). Pero, si leemos por columna o variable de arriba hacia abajo, tendríamos en la primera (género) dos hombres y dos mujeres (1, 1, 2, 2).

Por lo general, en la parte superior de la matriz de datos aparecen las opciones de los comandos para operar el programa de análisis estadístico como cualquier otro software (Archivo, Edición o Editar datos, etc.). Una vez que estamos seguros que no hay errores en la matriz, procedemos a realizar el análisis de la misma, el análisis estadístico. En cada programa tales opciones varían, pero en cuestiones mínimas.

Ahora, comentaremos brevemente los programas más importantes y de dos de ellos señalaremos sus comandos generales.

Statistical Package for the Social Sciences o Paquete Estadístico para las Ciencias Sociales (IBM® SPSS)

El SPSS (Paquete Estadístico para las Ciencias Sociales), desarrollado en la Universidad de Chicago, es uno de los más difundidos y actualmente es propiedad de IBM®. Contiene todos los análisis estadísticos que se describirán en este capítulo. En Iberoamérica, algunas instituciones educativas tienen versiones antiguas del SPSS; otras, versiones más recientes, ya sea en español o inglés y en los distintos sistemas operativos como Windows, Macintosh y UNIX.

Centro de recursos en línea

Como ocurre con otros programas, el IBM® SPSS se actualiza constantemente con versiones nuevas en varios idiomas.[1]

Asimismo, cada año surgen textos o manuales acordes con estas nuevas versiones. Sin embargo, en el centro de recursos el lector encontrará un manual que abarca las cuestiones esenciales de este paquete de análisis. Lo mejor para mantenerse al día en materia de SPSS® es consultar el sitio de IBM®que corresponda a su país (con las palabras clave: "IBM SPSS" puede encontrarlo mediante cualquier motor de búsqueda como Google). Asimismo, se puede "bajar" o "descargar" a la computadora una demostración del programa por tiempo limitado. Para la actualización de manuales, las palabras clave serían: "SPSS manuals" o "SPSS manuales".

La empresa IBM® afirma que se pueden solucionar diversos problemas de investigación con la suite de software IBM® SPSS Statistics, de la cual la versión "Base" contiene alrededor de 80% de los análisis. Para cuestiones más específicas se pueden adquirir diversos módulos que la compañía vende por separado, los cuales se integran a la versión *Base* con otros títulos o nombres.

Hasta agosto de 2013, las versiones más recientes tenían mejoras como mayor facilidad de uso, nuevas técnicas de análisis, mayor rendimiento y mejor integración con los demás productos IBM® (IBM SPSS®, 2013)

Como dijimos, el IBM® SPSS contiene las dos partes citadas que se denominan: *a*) vista de variables (para definiciones de las variables y consecuentemente, de los datos) y *b*) vista de los datos (matriz de datos). En ambas vistas se observan los comandos para operar en la parte superior.

[1] En el otoño de 2013, la familia de IBM SPSS Statistics contaba con alrededor de 30 productos con diferentes grados de complejidad. A éstos habría que añadirles los productos de las familias Collaboration and Deployment Services, Analytical Decision Management, Data Collection y Modeler.

El paquete IBM® SPSS trabaja de una manera muy sencilla: éste abre la matriz de datos y el investigador usuario selecciona las opciones más apropiadas para su análisis, tal como se hace en otros programas.

A continuación se describen de manera general algunas funciones principales del programa. Para profundizar y aprender su uso recomendamos revisar el manual de IBM SPSS®, que puede encontrar en el centro de recursos en línea (en el apartado de "Manuales auxiliares").

File (archivo): este botón tiene las siguientes funciones:

- Nuevo. Sirve para construir un archivo nuevo.
- Abrir. Se utiliza para abrir un archivo de datos previamente generado, de sintaxis, resultados o de proceso.
- Abrir base de datos. Tiene la función de generar, editar y ejecutar consultas en bases de datos previamente creadas.
- Leer datos de texto. Se usa para abrir archivos de texto.
- Cerrar. Como es común, sirve para cerrar el archivo en el que se está trabajando.
- Guardar. De igual forma, su función es guardar el archivo que se encuentra en uso.
- Guardar como. Se utiliza para guardar el archivo en el que se está trabajando con un nombre distinto al que ya tiene.
- Mostrar información del archivo de datos. Se trata de un archivo de trabajo o archivo externo.
- Caché de los datos. Es una copia temporal de los datos, la cual ayuda a mejorar el rendimiento cuando los archivos grandes de datos se leen desde una fuente externa. Si bien el archivo activo virtual puede ayudar a disminuir la cantidad de espacio en disco temporal, el no tener una copia temporal del archivo en activo provoca que la fuente de datos original se tenga que leer en cada procedimiento.
- Repositorio. Sirve para conectar, almacenar desde SPSS Statistics, publicar en la web, añadir un archivo, recuperar en SPSS Statistics o descargar un archivo.
- Presentación preliminar. Muestra en pantalla completa la tarea que se está llevando a cabo.
- Imprimir. Se utiliza para imprimir la tarea actual.
- Datos usados recientemente. Indica los datos que se ocuparon recientemente.
- Archivos usados recientemente. Muestra los últimos archivos utilizados.
- Salir. Cierra el programa IBM SPSS®.

Edit (edición): se emplea para modificar archivos, manipular la matriz (deshacer y rehacer acciones; cortar, copiar y pegar datos; pegar, borrar e insertar variables o casos), buscar y reemplazar datos e ir a casos o variables particulares entre otras acciones de edición.

View (ver): como su nombre lo dice es para visualizar la barra de estado, editor de menús, fuentes, cuadrícula, etiquetas de valor, marcar datos imputados, personalizar la vista de variables, etcétera.

Data (datos): en este menú se pueden definir las propiedades de las variables así como el nivel de medición, copiar propiedades de datos, establecer un nuevo atributo personalizado, definir fechas y conjuntos de respuestas múltiples, validación, identificación de casos duplicados y atípicos, ordenar casos y variables, trasponer, fundir archivos, reestructurar, agregar, diseño ortogonal, copiar conjunto de datos, segmentar archivos, seleccionar y ponderar casos.

Transform (transformar): con este botón se despliegan las opciones de calcular variable (crear variables compuestas por varios ítems o indicadores), contar valores dentro de los casos, valores de cambio, recodificar en las mismas o en distintas variables de manera personalizada o automática, agrupación visual, intervalos óptimos, preparar datos para modelado, asignar rangos a casos, asistente para fecha y hora, crear serie temporal, reemplazar valores perdidos y generar números aleatorios.

Analyze (analizar): por medio de esta opción se pueden solicitar análisis estadísticos que básicamente serían:

1. Informes (resúmenes de casos, información de columnas y renglones)
2. Estadísticos descriptivos (tablas de frecuencias, medidas de tendencia central y dispersión, razones, tablas de contingencia)

www.elosopanda.com | jamespoetrodriguez.com

3. Tablas (personalización de las tablas)
4. Comparar medias (prueba *t* y análisis de varianza —ANOVA— unidireccional)
5. Modelo lineal general (independiente o factor y dependiente, con covariable)
6. Modelos lineales generalizados
7. Modelos mixtos
8. Correlaciones (bivariada —dos— y multivariadas —tres o más—) para cualquier nivel de medición de las variables
9. Regresión (lineal, curvilineal y múltiple)
10. Loglineal
11. Redes neuronales
12. Clasificación (conglomerados y análisis discriminante)
13. Reducción de dimensiones (análisis de factores)
14. Escala (fiabilidad y escalamiento multidimensional)
15. Pruebas no paramétricas
16. Predicciones
17. Supervivencia
18. Respuesta múltiple (escalas)
19. Análisis de valores perdidos
20. Imputación múltiple
21. Muestras complejas
22. Control de calidad
23. Curva COR

Direct marketing (marketing directo): mediante esta nueva función se pueden clasificar y agrupar los datos de, en el caso de las empresas, sus clientes para obtener una comprensión más profunda de éstos.

Graphs (gráficos): con esta función se solicitan gráficos (barras en formato unidimensional y 3D, líneas, áreas, de sectores o pastel, máximos y mínimos, diagramas de caja, barras de error, pirámide de población, dispersión, histograma, etcétera).

Utilities (utilidades o herramientas): se definen ambientes, conjuntos, información sobre variables, etcétera.

Window (ventana): sirve para moverse a través de archivos y hacia otros programas.

Help (ayuda): cuenta con contenidos de ayuda, cómo utilizar SPSS, comandos, guías, "asesor estadístico" y demás elementos aplicados al paquete (con índice).

Minitab®

Minitab es un paquete que goza de popularidad por su relativo bajo costo. Incluye un considerable número de pruebas estadísticas y cuenta con un tutorial para aprender a utilizarlo y practicar; además, es muy sencillo de manejar.

Minitab tiene un sitio web (http://www.minitab.com/) en el cual se puede descargar una versión de prueba gratuita por tiempo limitado.

Para comenzar a utilizar Minitab, se abre una sesión (la cual se define con nombre y fecha) y se abre una matriz u hoja de trabajo (en la parte superior de la pantalla aparece la sesión y en la parte inferior se presenta la matriz). Se definen las variables (*C* —columnas—): nombre, formato (numérico, texto, fecha/tiempo), ancho (en dígitos), su descripción y orden de los valores. Los renglones o filas son casos. Los análisis realizados aparecen en la sesión (parte o pantalla superior) y las gráficas se reproducen en recuadros.

Entre sus comandos están los siguientes:

Archivo: sirve para construir un nuevo archivo, localizar uno ya construido, guardar o abrir archivos, abrir una gráfica de Minitab, especificar impresora, imprimir, cerrar, entre otras funciones.

Editar: útil para modificar archivos, buscar datos, copiar, cortar y eliminar celdas, conectar Minitab con otras aplicaciones, etcétera.

Datos: Se utiliza para ajustar o combinar columnas, incluye dividir la matriz, copiar o eliminar columnas y renglones o filas, establecer rangos, recodificar, cambiar el tipo de datos, desplegar datos, mostrar los datos de la hoja de trabajo en la ventana de sesión, entre otros.

Calcular: calcula las estadísticas de columnas y filas, distribuciones de probabilidad, matrices, estandarizaciones, operaciones aritméticas, etcétera.

Estadísticas: de manera fundamental, ejecuta los siguientes tipos de estadísticas:

1. Básicas: descriptiva e inferencial como distribución normal, prueba *t*, prueba de hipótesis acerca de la media poblacional, correlación, covarianza y *Chi* cuadrada.
2. Regresión lineal y múltiple.
3. Análisis de varianza (ANOVA) unidireccional y factorial.
4. DOE (análisis para diseños experimentales, análisis de respuestas).
5. Gráficas de control: de atributos, multivariados, de tiempo, individuales y grupales.
6. Herramientas de calidad: diagramas de dispersión, Pareto, causa-efecto, entre otros.
7. Confiabilidad: análisis de distribución, planes de prueba, análisis de garantía, prueba acelerada de vida útil, etcétera.
8. Análisis multivariado: conglomerados, análisis de factores (validación), análisis discriminante, análisis de conglomerados, de correspondencia simple o múltiple.
9. Series de tiempos: autocorrelación, correlación parcial, correlación cruzada, entre otras.
10. Tablas: tabulación cruzada, C*hi* cuadrada.
11. Estadística no paramétrica.
12. EDA (análisis exploratorio de datos, diagramas de caja, fotograma, etcétera).
13. Poder y tamaño de muestra (1-muestra *z*, 1-muestra-*t*, 2-muestra-*t*, ANOVA y otras. Sirve para determinar si el tamaño de muestra es apropiado para varias pruebas estadísticas).

Gráfica: sirve para solicitar gráficos, histogramas, barras de pastel, diagramas de dispersión, Pareto, series de tiempos, etcétera.

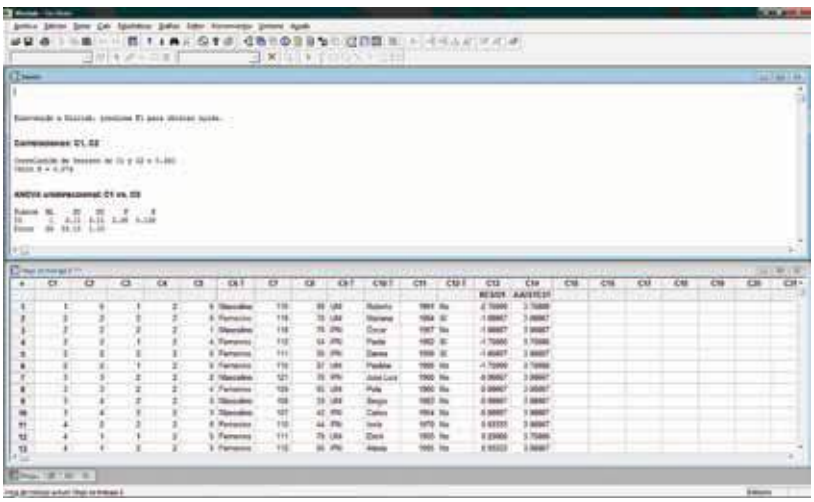
Editor: Se usa para mover, redefinir e insertar columnas, buscar o ir a un caso particular, entre otras acciones.

Herramientas: es útil para definir ambientes, conjuntos, información sobre variables, conexión a internet, consultas, etcétera.

Ventana: sirve para moverse a través de archivos y hacia otros programas, minimizar ventanas y demás funciones similares a otros programas.

Ayuda: cuenta con contenidos de ayuda, cómo utilizar Minitab, comandos, guías y demás elementos de Windows aplicados al paquete. En la figura 10.2 se muestra una vista de la pantalla de Minitab.

Figura 10.2 Pantalla de Minitab.



Otro programa de análisis sumamente difundido es el SAS (Sistema de Análisis Estadístico), que fue diseñado en la Universidad de Carolina del Norte. Es muy poderoso y su utilización se ha incrementado notablemente. Es un paquete muy completo para computadoras personales que contiene una variedad considerable de pruebas estadísticas (análisis de varianza, regresión, análisis de datos categóricos, análisis no paramétricos, etc.). Su página es: http://www.sas.com/technologies/analytics/statistics/stat/.

Centro de recursos en línea

En el centro de recursos en línea se incluye un vínculo al sitio de Decision Analyst, donde se puede descargar una versión de prueba del software STATS®, el cual ha sido incluido desde la segunda edición de esta obra y contiene los análisis bivariados más elementales para comenzar a comprenderlos y practicarlos. Por otro lado, en internet se encuentran diversos programas gratuitos de análisis estadístico para cualquier ciencia o disciplina.

Por lo general se elige el programa de análisis que está disponible en nuestra institución educativa, centro de investigación u organización de trabajo, o el que podamos comprar u obtener en internet. Todos los programas mencionados son excelentes opciones. Cualquiera nos sirve, solamente que debemos seleccionar uno. Recomendamos que en el centro de cómputo de su institución soliciten información respecto de los programas disponibles.

Paso 2: ejecutar el programa

La mayoría de los programas son fáciles de usar, pues lo único que hay que hacer es solicitar los análisis requeridos seleccionando las opciones apropiadas.

Paso 3: explorar los datos

En esta etapa, inmediata a la ejecución del programa, se inicia el análisis. Cabe señalar que si hemos llevado a cabo la investigación reflexionando paso a paso, la fase analítica es relativamente sencilla, porque: 1) formulamos las preguntas de investigación que pretendemos contestar, 2) visualizamos un alcance (exploratorio, descriptivo, correlacional o explicativo), 3) establecimos nuestras hipótesis (o estamos conscientes de que no las tenemos), 4) definimos las variables, 5) elaboramos un instrumento (conocemos qué ítems o indicadores miden qué variables y qué nivel de medición tiene cada variable: nominal, ordinal, de intervalos o razón) y 6) recolectamos los datos. Sabemos qué deseamos hacer, es decir, tenemos claridad.

La exploración típica se muestra en la figura 10.3 (que se hizo con base en el programa SPSS, pues, insistimos, puede variar de programa a programa en cuanto a comandos o instrucciones, pero no en lo referente a las funciones implementadas). Algunos conceptos pueden, por ahora, no significar nada para el lector que se inicia en los menesteres de la investigación, pero se irán explicando a lo largo del capítulo.

Veamos ahora los conceptos estadísticos que se aplican a la exploración de datos, pero antes es necesario realizar un par de apuntes, uno sobre las *variables del estudio* y las *variables de la matriz de datos*, y el otro sobre los factores de los que depende el análisis.

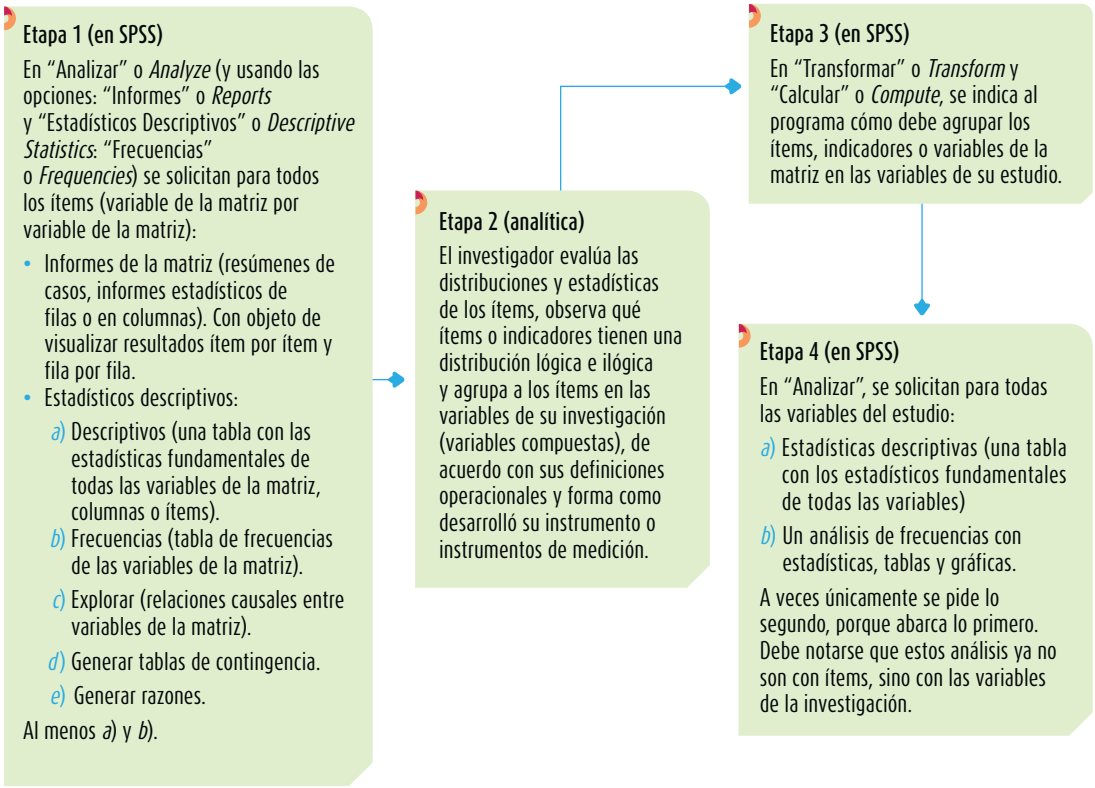
Apunte 1

Al final del capítulo anterior se introdujo el concepto de **variable de la matriz de datos**, que es distinto del concepto **variable de la investigación**. Las variables de la matriz de datos son columnas que constituyen indicadores o ítems. Las variables de la investigación son las propiedades medidas y que forman parte de las hipótesis o que se pretenden describir (género, edad, actitud hacia el presidente municipal, inteligencia, duración de un material, presión arterial, etc.). En ocasiones, las *variables de la investigación* requieren un único ítem, lectura o indicador para ser medidas (como en la tabla 10.2 con la variable "tipo de escuela a la que asiste"), pero en otras se necesitan varios ítems para tal finalidad. Cuando sólo se precisa de un ítem o indicador, las variables de la investigación

Variables de la matriz de datos Son columnas constituidas por ítems.

Variables de la investigación Son las propiedades medidas y que forman parte de las hipótesis o que se pretenden describir.

Figura 10.3 Secuencia más común para explorar datos en SPSS.



ocupan una columna de la matriz (una variable de la matriz). Pero si están compuestas de varios ítems, ocuparán tantas columnas como ítems (o variables en la matriz) las conformen. Esto se ejemplifica en la tabla 10.2 con los casos de las variables "presión arterial", "satisfacción respecto al superior" y "moral de los empleados".

Y cuando las *variables de la investigación* se integran de varios ítems o variables en la matriz, las columnas pueden ser continuas o no (estar ubicadas de manera seguida o en distintas partes de la matriz). En el cuarto ejemplo (variable "moral de los empleados"), las preguntas podrían ser las número 1, 2, 3, 4 y 5 del cuestionario; entonces, las primeras cinco columnas de la matriz representarán a estos ítems. Pero pueden ubicarse en distintos segmentos del cuestionario (por ejemplo, ser las preguntas 1, 5, 17, 22 y 38); entonces, las columnas que las representen se ubicarán de forma discontinua (serán las columnas o variables de la matriz 1, 5, 17, 22 y 38); porque regularmente la secuencia de las columnas corresponde a la secuencia de los ítems en el instrumento de medición.

Esta explicación se hace porque hemos visto estudiantes que confunden las variables de la matriz de datos con las variables del estudio. Son cuestiones vinculadas pero distintas.

Cuando una variable de la investigación está integrada por diversas variables de la matriz o ítems, suele llamarse *variable compuesta* y su puntuación total es el resultado de adicionar los valores de los reactivos que la conforman. Tal vez el caso más claro es la escala de Likert, en la que se suman las puntuaciones de cada ítem y se logra la calificación final. A veces la adición es una sumatoria, otras ocasiones es multiplicativa, un promedio o de otras formas, según se haya desarrollado el instrumento. Al ejecutar el programa y durante la fase exploratoria, se toman en cuenta todas las *variables de la investigación* e *ítems* y se considera a las *variables compuestas*, entonces se indica en el programa cómo están constituidas, mediante algunas instrucciones (en cada programa son distintas en cuanto al nombre, pero su función es similar). Por ejemplo, en SPSS se crean nuevas variables compuestas en la matriz de datos con el comando "Transformar" y luego con el comando "Calcular variable", de este modo, se construye la variable compuesta mediante una expresión numérica. Revisemos un ejemplo.

Tabla 10.2 Ejemplos de variables de investigación y formulación de ítems

Variable: tipo de escuela a la que asiste (con un ítem)	Variable: presión arterial (con dos indicadores)	Variable: satisfacción respecto al superior (con tres ítems)	Variable: moral del departamento donde se trabaja (con cinco ítems)
¿Asiste a una escuela pública o privada? [1] Escuela pública [2] Escuela privada	Lectura de la presión arterial sistólica:	1. ¿En qué medida está usted satisfecho con su superior inmediato? [1] Sumamente insatisfecho [2] Más bien insatisfecho [3] Ni insatisfecho ni satisfecho [4] Más bien satisfecho [5] Sumamente satisfecho	1. "En el departamento donde trabajo nos mantenemos unidos". [5] Totalmente de acuerdo [4] De acuerdo [3] Ni de acuerdo ni en desacuerdo [2] En desacuerdo [1] Totalmente en desacuerdo
	Lectura de la presión arterial diastólica:	2. ¿Qué tan satisfecho está usted con el trato que recibe de parte de su superior inmediato? [1] Sumamente insatisfecho [2] Más bien insatisfecho [3] Ni insatisfecho ni satisfecho [4] Más bien satisfecho [5] Sumamente satisfecho	2. "La mayoría de las veces en mi departamento compartimos la información más que guardarla para nosotros". [5] Totalmente de acuerdo [4] De acuerdo [3] Ni de acuerdo ni en desacuerdo [2] En desacuerdo [1] Totalmente en desacuerdo
		3. ¿Qué tan satisfecho está con la orientación que le proporciona su superior inmediato para que usted realice su trabajo? [1] Sumamente insatisfecho [2] Más bien insatisfecho [3] Ni insatisfecho ni satisfecho [4] Más bien satisfecho [5] Sumamente satisfecho	3. "En mi departamento nos mantenemos en contacto permanentemente". [5] Totalmente de acuerdo [4] De acuerdo [3] Ni de acuerdo ni en desacuerdo [2] En desacuerdo [1] Totalmente en desacuerdo
			4. "En mi departamento nos reunimos con frecuencia para hablar tanto de asuntos de trabajo como de cuestiones personales". [5] Totalmente de acuerdo [4] De acuerdo [3] Ni de acuerdo ni en desacuerdo [2] En desacuerdo [1] Totalmente en desacuerdo
			5. "En mi trabajo todos nos llevamos bien". [5] Totalmente de acuerdo [4] De acuerdo [3] Ni de acuerdo ni en desacuerdo [2] En desacuerdo [1] Totalmente en desacuerdo
Esta variable es medida por una sola pregunta y ocupa una columna o variable de la matriz.	Esta variable es medida por dos indicadores o lecturas y ocupa dos columnas o variables de la matriz.	Esta variable es medida por tres preguntas y ocupa tres columnas o variables de la matriz.	Esta variable es medida por cinco preguntas y ocupa cinco columnas o variables de la matriz.

En el caso de la variable "moral en el departamento donde se trabaja", podríamos asignar las siguientes columnas (en el supuesto de que fueran continuas) a los cinco ítems, tal como se muestra en la tabla 10.3.

Y tener la siguiente matriz (ejemplo):

Ejemplo

Tabla 10.3 Ejemplo con la variable moral

Variable de la investigación: moral	Variable de la matriz que corresponde a la variable de la investigación	Ubicación en la matriz
1. "En el departamento donde trabajo nos mantenemos unidos" 5 Totalmente de acuerdo 4 De acuerdo 3 Ni de acuerdo ni en desacuerdo 2 En desacuerdo 1 Totalmente en desacuerdo	Frase 1 (fr1)	Columna 1
2. "La mayoría de las veces en mi departamento compartimos la información más que guardarla para nosotros" 5 Totalmente de acuerdo 4 De acuerdo 3 Ni de acuerdo ni en desacuerdo 2 En desacuerdo 1 Totalmente en desacuerdo	Frase 2 (fr2)	Columna 2
3. "En mi departamento nos mantenemos en contacto permanentemente" 5 Totalmente de acuerdo 4 De acuerdo 3 Ni de acuerdo ni en desacuerdo 2 En desacuerdo 1 Totalmente en desacuerdo	Frase 3 (fr3)	Columna 3
4. "En mi departamento nos reunimos con frecuencia para hablar tanto de asuntos de trabajo como de cuestiones personales" 5 Totalmente de acuerdo 4 De acuerdo 3 Ni de acuerdo ni en desacuerdo 2 En desacuerdo 1 Totalmente en desacuerdo	Frase 4 (fr4)	Columna 4
5. "En mi trabajo todos nos llevamos muy bien" 5 Totalmente de acuerdo 4 De acuerdo 3 Ni de acuerdo ni en desacuerdo 2 En desacuerdo 1 Totalmente en desacuerdo	Frase 5 (fr5)	Columna 5

Casos	fr1	fr2	fr3	fr4	fr5
1	1	2	2	4	3
2	2	2	2	2	2
K	2	3	2	2	3

En las opciones "Transformar" y "Calcular" o "Computar" el programa nos pide que indiquemos el nombre de la nueva variable (en este caso la compuesta por cinco frases): *moral*. Y nos solicita que desarrollemos la expresión numérica que corresponda a esta variable compuesta: *fr1+fr2+fr3+fr4+fr5* (automáticamente el programa realiza la operación y agrega la nueva variable compuesta "moral" a la matriz de datos y realiza los cálculos, y ahora sí, la *variable del estudio* es una variable más de la matriz de datos). La matriz se modificaría de la siguiente manera:

Ejemplo

Casos	fr1	fr2	fr3	fr4	fr5	Moral
1	1	2	2	4	3	12
2	2	2	2	2	2	10
K	2	3	2	2	3	12

Desde luego, para mantener esta variable debemos demostrar que fue medida de forma confiable y válida, así como evaluar si todos los ítems aportan favorablemente a ambos elementos o algunos no. Y en lugar de una suma, la variable *moral* podría ser un promedio de las cinco frases o variables de la matriz (como ya se mencionó en el tema de la escala de Likert). Entonces, la expresión en "Calcular" es: (*fr1+fr2+fr3+fr4+fr5*)/*5*, y los valores en "moral" serían:

Ejemplo

Casos	fr1	fr2	fr3	fr4	fr5	Moral
1	1	2	2	4	3	2.4
2	2	2	2	2	2	2.0
K	2	3	2	2	3	2.4

Por último, las variables de la investigación son las que nos interesan, ya sea que estén compuestas por uno, dos, diez, 50 o más ítems. El primer análisis es sobre los ítems, únicamente para explorar; el análisis descriptivo final es sobre las *variables del estudio*.

Apunte 2

Los análisis de los datos dependen de tres factores:[2]

- ***a)*** El *nivel de medición* de las variables.
- ***b)*** La manera como se hayan formulado las *hipótesis*.
- ***c)*** El *interés analítico del investigador* (que depende del planteamiento del problema).

[2] Babbie (2014), O'Leary (2014), Hollander, Wolfe y Chicken (2013); Jarman (2013), Kon y Rai (2013), Hernández-Sampieri *et al.* (2013), Collier, LaPorte y Seawright (2012); Martin y Bridgmon (2012), Garson (2012), Howell (2011), Mertens (2010), Gershkoff (2008), Scott y Albaum (2006), Reynolds (1984) y Hildebbrand, Laing y Rosenthal (1977).

Por ejemplo, los análisis que se aplican a una variable nominal son distintos a los de una variable por intervalos. Se sugiere repasar los niveles de medición vistos en el capítulo anterior.

El investigador busca, en primer término, describir sus datos y posteriormente efectuar análisis estadísticos para relacionar sus variables. Es decir, realiza análisis de estadística descriptiva para cada una de las variables de la matriz (ítems o indicadores) y luego para cada una de las variables del estudio, finalmente aplica cálculos estadísticos para probar sus hipótesis. Los tipos o métodos de análisis cuantitativo o estadístico son variados y se comentarán a continuación; pero cabe señalar que el análisis no es indiscriminado, sino que cada método tiene su razón de ser y un propósito específico; por ello, no deben hacerse más análisis de los necesarios. La estadística no es un fin en sí misma, sino una herramienta para evaluar los datos.

Estadística descriptiva para cada variable

2 La primera tarea es describir los datos, los valores o las puntuaciones obtenidas para cada variable. Por ejemplo, si aplicamos a 2 112 niños el cuestionario sobre los usos y las gratificaciones que la televisión tiene para ellos, ¿cómo pueden describirse estos datos? Esto se logra al describir la distribución de las puntuaciones o frecuencias de cada variable.

¿Qué es una distribución de frecuencias?

Distribución de frecuencias Conjunto de puntuaciones de una variable ordenadas en sus respectivas categorías.

Una **distribución de frecuencias** es un conjunto de puntuaciones respecto de una variable ordenadas en sus respectivas categorías y generalmente se presenta como una tabla (O'Leary, 2014 y Nicol, 2006).

La tabla 10.4 muestra un ejemplo de una distribución de frecuencias.

Ejemplo

En un estudio entre 200 personas latinas que viven en el estado de California, Estados Unidos,[3] se les preguntó: ¿cómo prefiere que se refieran a usted en cuanto a su origen étnico? Las respuestas fueron:

Tabla 10.4 Ejemplo de una distribución de frecuencias

Variable: preferencias al referir el origen étnico (nombrada en SPSS: prefoe)		
Categorías	**Códigos (valores)**	**Frecuencias**
Hispano	1	52
Latino	2	88
Latinoamericano	3	6
Americano	4	22
Otros	5	20
No respondieron	6	12
Total		200

A veces, las *categorías* de las distribuciones de frecuencias son tantas que es necesario resumirlas. Por ejemplo, examinemos detenidamente la distribución de la tabla 10.5. Esta distribución podría compendiarse como en la tabla 10.6.

[3] Encuesta con 7% de margen de error (University of Southern California y Bendixen and Associates, 2002).

Tabla 10.5 Ejemplo de una distribución que necesita resumirse

Variable: calificación en la prueba de motivación			
Categorías	Frecuencias	Categorías	Frecuencias
48	1	74	1
55	2	75	4
56	3	76	3
57	5	78	2
58	7	80	4
60	1	82	2
61	1	83	1
62	2	84	1
63	3	86	5
64	2	87	2
65	1	89	1
66	1	90	3
68	1	92	1
69	1	**Total**	63
73	2		

Tabla 10.6 Ejemplo de una distribución resumida

Variable: calificación en la prueba de motivación	
Categorías	Frecuencias
55 o menos	3
56-60	16
61-65	9
66-70	3
71-75	7
76-80	9
81-85	4
86-90	11
91-96	1
Total	63

¿Qué otros elementos contiene una distribución de frecuencias?

Las distribuciones de frecuencias pueden completarse agregando los porcentajes de casos en cada categoría, los porcentajes válidos (excluyendo los valores perdidos) y los porcentajes acumulados (porcentaje de lo que se va acumulando en cada categoría, desde la más baja hasta la más alta).

La tabla 10.7 muestra un ejemplo con las frecuencias y porcentajes en sí, los porcentajes válidos y los acumulados. El *porcentaje acumulado* constituye lo que aumenta en cada categoría de manera porcentual y progresiva (en orden de aparición de las categorías), tomando en cuenta los *porcentajes válidos*. En la categoría "sí se ha obtenido la cooperación", se ha acumulado 74.6%. En la categoría "no se ha obtenido la cooperación", se acumula 78.7% (74.6% de la categoría anterior y 4.1% de la categoría en cuestión). En la última categoría siempre se acumula el total (100%).[4]

Tabla 10.7 Ejemplo de una distribución de frecuencias con todos sus elementos

Variable: cooperación del personal con el proyecto de calidad de la empresa				
Categorías	Códigos	Frecuencias	Porcentaje válido	Porcentaje acumulado
Sí se ha obtenido la cooperación	1	91	74.6	74.6
No se ha obtenido la cooperación	2	5	4.1	78.7
No respondieron	3	26	21.3	100.0
Total		**122**	**100.0**	

Las columnas *porcentaje* y *porcentaje válido* son iguales (mismas cifras o valores) cuando *no* hay valores perdidos; pero si tenemos valores perdidos, la columna *porcentaje válido* presenta los cálculos sobre el total menos tales valores. En la tabla 10.8 se muestra un ejemplo con valores perdidos en el caso de un estudio exploratorio sobre los motivos de los niños celayenses para elegir su personaje televisivo favorito (García y Hernández-Sampieri, 2005).

[4] En variables nominales el porcentaje acumulado es relativo porque no hay orden o jerarquía entre categorías, pero se buscó un ejemplo simple para entender más fácilmente el concepto.

Al elaborar el informe de resultados, una distribución se presenta con los elementos más informativos para el lector y la descripción de los resultados o un comentario, tal como se muestra en la tabla 10.9.

Tabla 10.8 Ejemplo de tabla con valores perdidos (en SPSS)[5]

Motivos de la preferencia de su personaje favorito					
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Divertidos	142	72.1	73.2	73.2
	Buenos	10	5.1	5.2	78.4
	Tienen poderes	23	11.7	11.9	90.2
	Son fuertes	19	9.6	9.8	100.0
	Total	194	98.5	100.0	
Perdidos	No contestaron	3	1.5		
Total		**197**	**100.0**		

Tabla 10.9 Ejemplo de una distribución de frecuencias para presentar a un usuario

¿Se ha obtenido la cooperación del personal para el proyecto de calidad?		
Obtención	**Núm. de organizaciones**	**Porcentajes**
Sí	91	74.6
No	5	4.1
No respondieron	26	21.3
Total	122	100.0
COMENTARIO: Prácticamente tres cuartas partes de las organizaciones sí han obtenido la cooperación del personal. Llama la atención que poco más de una quinta parte no quiso comprometerse con su respuesta. Las empresas que no han logrado la cooperación del personal mencionaron como factores el absentismo, rechazo al cambio y conformismo.		

En los programas de análisis se solicita la distribución de frecuencias de cada variable de la investigación (por ejemplo, en SPSS: Analizar → Estadísticos descriptivos → Frecuencias).[6]

¿De qué otra manera pueden presentarse las distribuciones de frecuencias?

3 Las distribuciones de frecuencias, especialmente cuando utilizamos los porcentajes, pueden presentarse en forma de histogramas o gráficas de otro tipo (por ejemplo: de pastel). Algunos ejemplos se muestran en la figura 10.4.

SPSS, Minitab y SAS producen tales gráficas, o bien, los datos pueden exportarse a otros programas o paquetes que las generan (de cualquier tipo, a colores, utilizando efectos de movimiento y en tercera dimensión, como por ejemplo: Power Point).

Centro de recursos en línea

Para obtener las gráficas en SPSS no olvide consultar el centro de recursos en línea de esta obra el *manual* "Introducción al IBM SPSS®".

Polígonos de frecuencias Relacionan las puntuaciones con sus respectivas frecuencias por medio de gráficas útiles para describir los datos.

Las distribuciones de frecuencias también se pueden graficar como polígonos de frecuencias

Los **polígonos de frecuencias** relacionan las puntuaciones con sus respectivas frecuencias. Es más bien propio de un nivel de medición por intervalos o razón. Los

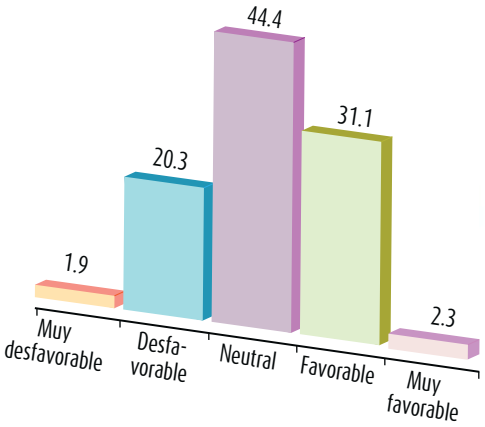
[5] En todos los casos, cifras redondeadas por el programa.

[6] Esta secuencia en SPSS para obtener los análisis de frecuencias requeridos, al igual que el resto de análisis (valores, tablas y gráficas), se incluyen en el manual "Introducción al IBM SPSS®", que puede descargarse del centro de recursos en línea.

Figura 10.4 Ejemplos de gráficas para presentar distribuciones.

Histogramas

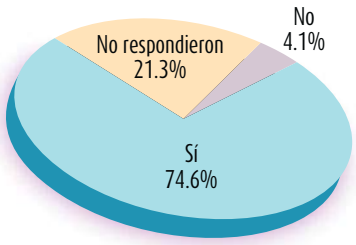
Opinión acerca del actual alcalde del municipio de San Martín Aurelio



Solamente la tercera parte de los ciudadanos expresa una opinión positiva respecto al alcalde (favorable o muy favorable).

Gráficas circulares

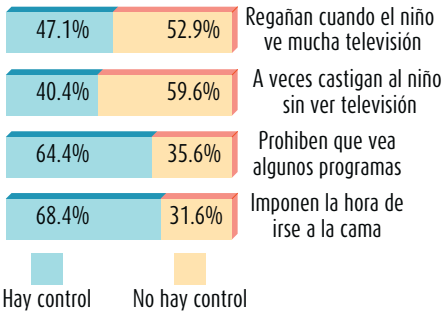
Cooperación de todo el personal (o la mayoría) para el proyecto de calidad (122 = 100%)



Prácticamente tres cuartas partes han obtenido la cooperación de todo el personal para el proyecto de la empresa. Pero llama la atención que poco más de una quinta parte no quiso comprometerse con su respuesta. Los cinco motivos de no cooperación con dicho proyecto fueron: absentismo, falta de interés, rechazo al cambio, falta de concientización y conformismo.

Otros tipos de gráficas

Control paterno sobre el uso que los niños hacen de la televisión.



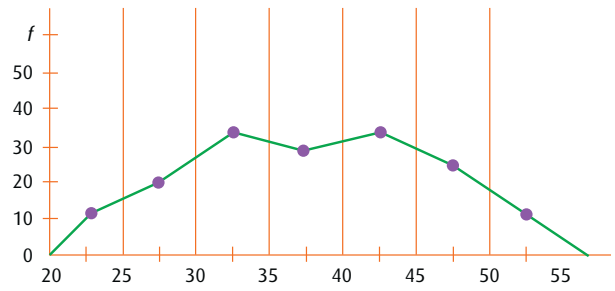
polígonos se construyen sobre los puntos medios de los intervalos. Por ejemplo, si los intervalos fueran 20-24, 25-29, 30-34, 35-39, y siguientes; los puntos medios serían 22, 27, 32, 37, etc. SPSS o Minitab realizan esta labor en forma automática. Un ejemplo de un polígono de frecuencias se muestra en la figura 10.5.

El polígono de frecuencias obedece a la siguiente distribución:

Categorías/intervalos	Frecuencias absolutas
20-24.9	10
25-29.9	20
30-34.9	35
35-39.9	33
40-44.9	36
45-49.9	27
50-54.9	8
TOTAL	169

Figura 10.5 Ejemplo de un polígono de frecuencias.

Variable: satisfacción en el trabajo



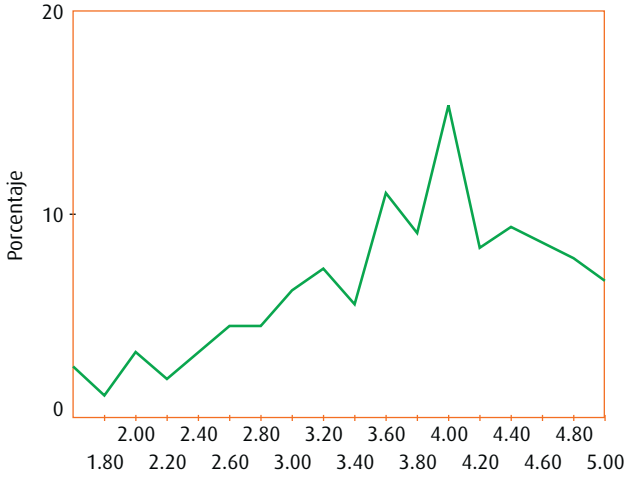
Los polígonos de frecuencias representan curvas útiles para describir los datos. Nos indican hacia dónde se concentran los casos (personas, organizaciones, segmentos de contenido, mediciones de polución, datos de presión arterial, etc.) en la escala de la variable; más adelante se hablará de ello.

En resumen, para cada una de las variables de la investigación se obtiene su distribución de frecuencias y se grafica (histograma, gráfica de barras, gráfica circular o polígono de frecuencias) (Huck, 2006).

En la figura 10.6 se muestra otro ejemplo.

Figura 10.6 Ejemplo de un polígono de frecuencias con la variable innovación.

Variable: innovación
Con respecto a la innovación en la empresa, que es la percepción del apoyo a las iniciativas tendientes a introducir mejoras en la manera como se realiza el trabajo, a nivel organizacional y departamental, la mayoría de los individuos tienden a estar en altos niveles de la escala.



El polígono puede presentarse con frecuencias como en la figura 10.5 o con porcentajes como con este segundo ejemplo. Pero además de la distribución o polígono de frecuencias, deben calcularse las *medidas de tendencia central* y de *variabilidad o dispersión.*

¿Cuáles son las medidas de tendencia central?

Medidas de tendencia central Valores medios o centrales de una distribución que sirven para ubicarla dentro de la escala de medición de la variable.

Las **medidas de tendencia central** son puntos en una distribución obtenida, los valores medios o centrales de ésta, y nos ayudan a ubicarla dentro de la escala de medición de la variable analizada. Las principales medidas de tendencia central son tres: *moda, mediana* y *media*. El nivel de medición de la variable determina cuál es la medida de tendencia central apropiada para interpretar (Graham, 2013, Kwok, 2008a y Platt, 2003a).

Moda Categoría o puntuación que se presenta con mayor frecuencia.

La **moda** es la categoría o puntuación que ocurre con mayor frecuencia. En la tabla 10.7, la moda es "1" (sí se ha obtenido la cooperación). Se utiliza con cualquier nivel de medición.

La mediana es el valor que divide la distribución por la mitad. Esto es, la mitad de los casos caen por debajo de la mediana y la otra mitad se ubica por encima de ésta. La mediana refleja la posición intermedia de la distribución (Hempel, 2006). Por ejemplo, si los datos obtenidos fueran:

24 31 35 35 38 43 45 50 57

La mediana es 38, porque deja cuatro casos por encima (43, 45, 50 y 57) y cuatro casos por debajo (35, 35, 31 y 24). Parte a la distribución en dos mitades. En general, para descubrir la puntuación que constituye la mediana de una distribución, simplemente se aplica la fórmula:

$$\frac{N+1}{2}$$

Si tenemos nueve casos, $\frac{9+1}{2}$ entonces buscamos el quinto valor y éste es la mediana. Note que la mediana es el valor observado que se localiza a la mitad de la distribución, no el valor de cinco. La fórmula no nos proporciona directamente el valor de la mediana, sino el número de caso en donde está la mediana.

La mediana es una medida de tendencia central propia de los niveles de medición ordinal, por intervalos y de razón. No tiene sentido con variables nominales, porque en este nivel no hay jerarquías ni noción de encima o debajo. Asimismo, la mediana es particularmente útil cuando hay valores extremos en la distribución. No es sensible a éstos. Si tuviéramos los siguientes datos:

24 31 35 35 38 43 45 50 248

la mediana seguiría siendo 38.

Para la interpretación de la media y la mediana, se incluye un comentario al respecto en el siguiente ejemplo.[7]

Ejemplo

¿Qué edad tiene? Si teme contestar no se preocupe, los perfiles de edad difieren de un país a otro.[8]

A mediados de 2013, la población mundial superó los 7 100 millones de habitantes y se espera que en 2015 seamos más de 7 300 millones de humanos y en 2050 pasemos de 9 000 millones (United States Census Bureau, 2013, Alberich, 2013 y Organización de las Naciones Unidas, 2011).

En 2012, la mediana de edad mundial fue de 29 años, lo que significa que la mitad de los habitantes del globo terrestre sobrepasa esta edad y el otro medio es más joven. Cabe señalar que la mediana varía de un lugar a otro. Por ejemplo, en 2012 por bloque de países fue de 28.1 años en Asia, 39.9 en Europa, 26.9 en Oceanía, 27.8 en Sudamérica, 30.2 en América Central y El Caribe (incluyendo México), 20.1 en África y 38.4 en Norteamérica (Canadá y Estados Unidos). Actualmente, países con población muy joven son Uganda y Níger, con una edad mediana de prácticamente 15 años, y entre los más viejos podemos incluir a Japón (44.8), Alemania (45.3) y Mónaco (49.9). En América Latina tenemos algunos ejemplos como: Perú (26.5), México (27.4), Colombia (28.3), Chile (32.8), República Dominicana (26.5), Brasil (29.6), Costa Rica (29.2), Panamá (27.8), Ecuador (26), Paraguay (25.9), Uruguay (33.8), Honduras (21.3), El Salvador (24.7), Argentina (30.7) y Bolivia (22.8).

La mediana de edad ha ido en constante ascenso durante el siglo pasado y el actual. Se estima que para 2050 la edad mediana mundial habrá aumentado a más de 36 años. Buena noticia para el actual ciudadano global medio, porque parece ser que se encuentra en la situación de "envejecer más lentamente".

Media Promedio aritmético de una distribución. Es la medida de tendencia central más utilizada.

La **media** es tal vez la medida de tendencia central más utilizada (Graham, 2013, Kwok, 2008b y Leech, Onwuegbuzie y Daniel, 2006) y puede definirse como el promedio aritmético de una distribución. Se simboliza como $\overline{X}$, y es la suma de todos los valores dividida entre el número de casos. Es una medida solamente aplicable a mediciones por intervalos o de razón. Carece de sentido para variables medidas en un nivel nominal u ordinal. Resulta sensible a valores extremos. Si tuviéramos las siguientes puntuaciones:

8 7 6 4 3 2 6 9 8

El promedio sería igual a 5.88. Pero bastaría una puntuación extrema para alterarlo de manera notoria:

8 7 6 4 3 2 6 9 **20** (promedio igual a 7.22).

La mediana puede ser una medida de interpretación más útil que la media si la distribución está más cargada hacia puntuaciones extremas (Kwok, 2008a y Hempel, 2006).

El cálculo de la media lo podrá encontrar el lector en el centro de recursos en línea de la obra: Material complementario → Documentos → Documento 2, "Fórmulas y procedimientos estadísticos".

Centro de recursos en línea

¿Cuáles son las medidas de la variabilidad?

Medidas de la variabilidad Intervalos que indican la dispersión de los datos en la escala de medición de la variable.

Las **medidas de la variabilidad** indican la dispersión de los datos en la escala de medición de la variable considerada y responden a la pregunta: ¿dónde están diseminadas las puntuaciones o los valores obtenidos? Las medidas de tendencia central son valores en una distribución y las medidas de la variabilidad son intervalos que designan distancias o un número de unidades en la escala de medición (Kon y Rai, 2013 y O'Brien, 2007). Las medidas de la variabilidad más utilizadas son *rango*, *desviación estándar* y *varianza*.

[7] Basado en una idea de Leguizamo (1987).

[8] Datos obtenidos de Getamap (2013), Kaiser Family Foundation (2013) y Worldstat (2013). Son estimaciones con un margen de error de 1% y se basan en información de 2010-2011 y proyecciones a 2012.

Rango Extensión total de los datos en la escala.

El **rango**, también llamado *recorrido*, es la diferencia entre la puntuación mayor y la puntuación menor, e indica el número de unidades en la escala de medición que se necesitan para incluir los valores máximo y mínimo. Se calcula así: $X_M - X_m$ (puntuación mayor menos puntuación menor). Si tenemos los siguientes valores:

17 18 20 20 24 28 28 30 33

El rango será: 33 – 17 = 16.

Cuanto *más grande* sea el *rango*, *mayor* será la *dispersión de los datos* de una distribución.

Desviación estándar Promedio de desviación de las puntuaciones con respecto a la media que se expresa en las unidades originales de medición de la distribución.

La **desviación estándar** o característica es el promedio de desviación de las puntuaciones con respecto a la media (Jarman, 2013 y Levin, 2003). Esta medida se expresa en las unidades originales de medición de la distribución. Se interpreta en relación con la media. Cuanto mayor sea la dispersión de los datos alrededor de la media, mayor será la desviación estándar. Se simboliza como: s o la sigma minúscula σ, o bien mediante la abreviatura DE. Su cálculo lo podrá encontrar el lector en el centro de recursos, en: Material complementario → Documentos → Documento 2, "Fórmulas y procedimientos estadísticos".

Centro de recursos en línea

La desviación estándar se interpreta como *cuánto se desvía, en promedio, de la media un conjunto de puntuaciones*.

Supongamos que un investigador obtuvo para su muestra una media (promedio) de ingreso familiar anual de 6 000 unidades monetarias y una desviación estándar de 1 000. La interpretación es que los ingresos familiares de la muestra se desvían, en promedio, mil unidades monetarias respecto a la media.

La desviación estándar sólo se utiliza en variables medidas por intervalos o de razón.

La varianza

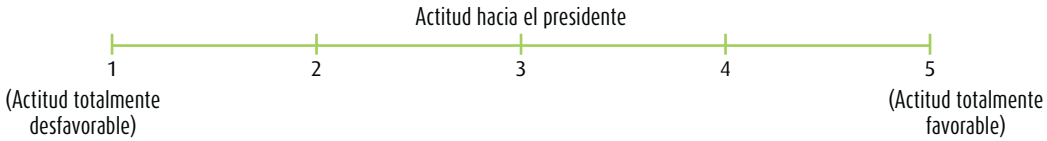
Varianza Desviación estándar elevada al cuadrado.

La **varianza** es la desviación estándar elevada al cuadrado y se simboliza como s^2. Es un concepto estadístico muy importante, ya que la mayoría de las pruebas cuantitativas se fundamentan en él. Diversos métodos estadísticos parten de la descomposición de la varianza (Zhang, 2013; Beins y McCarthy, 2009; Wilcox, 2008; y Jackson, 2008). Sin embargo, con fines descriptivos se utiliza preferentemente la desviación estándar.

¿Cómo se interpretan las medidas de tendencia central y de la variabilidad?

Cabe destacar que al describir nuestros datos, respecto a cada *variable del estudio*, interpretamos las medidas de tendencia central y de la variabilidad en conjunto, no aisladamente. Consideramos todos los valores. Para interpretarlos, lo primero que hacemos es tomar en cuenta el rango potencial de la escala. Supongamos que aplicamos una escala de actitudes del tipo Likert para medir la "actitud hacia el presidente" de una nación (digamos que la escala tuviera 18 ítems y se promediaran sus valores). El rango potencial es de uno a cinco (véase la figura 10.7).

Figura 10.7 Ejemplo de escala con rango potencial.



Si obtuviéramos los siguientes resultados:

Variable: actitud hacia el presidente
Moda: 4.0

Mediana: 3.9
Media ($\overline{X}$): 4.2
Desviación estándar: 0.7
Puntuación más alta observada (máximo): 5.0
Puntuación más baja observada (mínimo): 2.0
Rango: 3

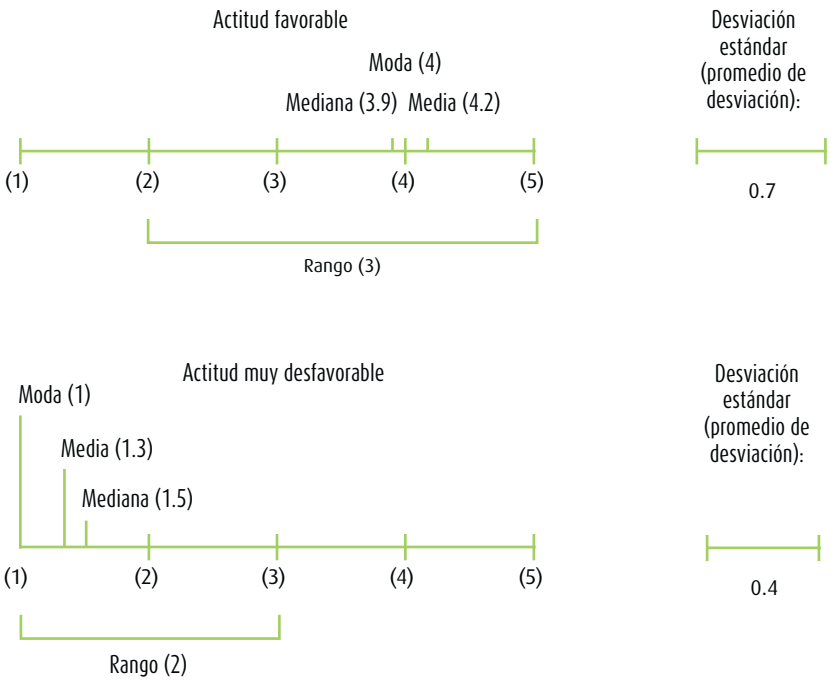
podríamos hacer la siguiente interpretación descriptiva: la actitud hacia el presidente es favorable. La categoría que más se repitió fue 4 (favorable). Cincuenta por ciento de los individuos está por encima del valor 3.9 y el restante 50% se sitúa por debajo de este valor (mediana). En promedio, los participantes se ubican en 4.2 (favorable). Asimismo, se desvían de 4.2, en promedio, 0.7 unidades de la escala. Ninguna persona calificó al presidente de manera muy desfavorable (no hay "1"). Las puntuaciones tienden a ubicarse en valores medios o elevados.

En cambio, si los resultados fueran:

Variable: actitud hacia el presidente
Moda: 1
Mediana: 1.5
Media ($\overline{X}$): 1.3
Desviación estándar: 0.4
Máximo: 3.0
Mínimo: 1.0
Rango: 2.0

La interpretación es que la actitud hacia el presidente es muy desfavorable. En la figura 10.8 vemos gráficamente la comparación de resultados. La variabilidad también es menor en el caso de la actitud muy desfavorable (los datos se encuentran menos dispersos).

Figura 10.8 Ejemplo de interpretación gráfica de las estadísticas descriptivas.



Otro ejemplo de interpretación de los resultados de una medición respecto a una variable es el que ahora se presenta.

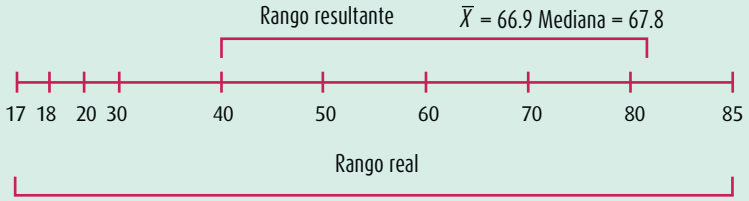
Ejemplo

Hernández-Sampieri y Cortés (1982) aplicaron una prueba de motivación intrínseca sobre la ejecución de una tarea a 60 participantes de un experimento. La escala contenía 17 ítems (con cinco opciones cada uno, uno a cinco) y los resultados fueron los siguientes:[9]

n: 60	Rango: 41	Mínimo: 40	Máximo: 81
Media: 66.9	Mediana: 67.8	Moda: 61	DE: 9.1
Varianza: 83	Curtosis: 0.6	Asimetría: −0.8	EE: 1.18
Sumatoria: 4 013			

¿Qué podríamos decir sobre la motivación intrínseca de los participantes?

El nivel de motivación intrínseca exhibido por los participantes tiende a ser elevado, como lo indican los resultados. El rango real de la escala iba de 17 a 85. El rango resultante para esta investigación varió de 40 a 81. Por tanto, es evidente que los individuos se inclinaron hacia valores elevados en la medida de motivación intrínseca. Además, la media de los participantes es de 66.9 y la mediana de 67.8, lo cual confirma la tendencia de la muestra hacia valores altos de la escala. A pesar de que la dispersión de las puntuaciones de los sujetos es considerable (la desviación estándar es igual a 9.1 y el rango es de 41), esta dispersión se manifiesta en el área más elevada de la escala. Veámoslo gráficamente.



Escala de motivación intrínseca (datos ordinales, supuestos como datos en nivel de intervalo).

En resumen, la tarea resultó intrínsecamente motivante para la mayoría de los participantes; sólo que para algunos resultó muy motivante, para otros, relativamente motivante, y para los demás, medianamente motivante. Esto es, que la tendencia general es hacia valores superiores.

Ahora bien, ¿qué significa un alto nivel de motivación intrínseca exhibido con respecto a una tarea? Implica que la tarea fue percibida como atractiva, interesante, divertida y categorizada como una experiencia agradable. Asimismo, involucra que los individuos, al ejecutarla, derivaron de ella sentimientos de satisfacción, goce y realización personal. Por lo general, quien se encuentra intrínsecamente motivado hacia una labor, disfrutará la ejecución de ésta, ya que obtendrá de la labor *per se* recompensas internas, como sentimientos de logro y autorrealización. Además de ser absorbido por el desarrollo de la tarea y, al tener un buen desempeño, la opinión de sí mismo mejorará o se verá reforzada.

¿Hay alguna otra estadística descriptiva?

Sí, *la asimetría y la curtosis.* Los *polígonos de frecuencia* son *curvas*, por ello se representan como tales (figura 10.9), para que puedan analizarse en términos de probabilidad y visualizar su grado de dispersión. Estos dos elementos resultan esenciales para analizar estas curvas o polígonos de frecuencias.

Asimetría y **curtosis** Estadísticas que se usan para conocer cuánto se parece una distribución a la distribución teórica llamada *curva normal* o campana de Gauss y dónde se concentran las puntuaciones.

La **asimetría** es una estadística necesaria para conocer cuánto se parece nuestra distribución a una distribución teórica llamada *curva normal* (la cual se representa también en la figura 10.9) y constituye un indicador del lado de la curva donde se agrupan las frecuencias. Si es cero (asimetría = 0), la curva o distribución es simétrica. Cuando es positiva, quiere decir que hay más valores agrupados hacia la izquierda de la curva (por debajo de la media). Cuando es negativa, significa que los valores tienden a agruparse hacia la derecha de la curva (por encima de la media) (Hume, 2011; Taylor, 2007a; Salkind, 2006; y Burkhart, 2003).

La **curtosis** es un indicador de lo plana o "picuda" que es una curva. Cuando es cero (curtosis = 0), significa que puede tratarse de una *curva normal*. Si es positiva, quiere decir que la curva, la distri-

[9] EE significa "error estándar".

- Observación durante la inmersión en la comunidad (contexto).
- Entrevistas con mujeres que la padecen.
- Entrevistas con sus familiares.
- Observación inmediatamente posterior al parto (durante la convalecencia) en hospitales rurales o en sus hogares (en diversas comunidades indígenas, las mujeres paren en su casa).
- Algún grupo de enfoque con mujeres que la han experimentado.

Triangulación de datos Utilización de diferentes fuentes y métodos de recolección.

De esta manera, el sentido de comprender la depresión posparto en tal comunidad será mayor que si únicamente llevamos a cabo entrevistas.

Al hecho de utilizar diferentes fuentes y métodos de recolección se le denomina **triangulación de datos**. Sobre este tema regresaremos de manera recurrente.

El análisis de los datos cualitativos

En el proceso cuantitativo primero se recolectan todos los datos y luego se analizan, mientras que en la investigación cualitativa no es así, sino que la recolección y el análisis ocurren prácticamente en paralelo; además, el análisis no es uniforme, ya que cada estudio requiere un esquema peculiar. Sin embargo, diversos autores hemos propuesto un análisis genérico o básico común a diversas investigaciones cualitativas, porque en nuestra experiencia, a la mayoría de los alumnos que se inician en la indagación cualitativa —sobre todo si el único enfoque que han utilizado es el cuantitativo— les cuesta trabajo comenzar su tarea analítica y se preguntan por dónde empezar y qué secuencia establecer.

En este apartado se propone un proceso de análisis que incorpora las concepciones de diversos teóricos de la metodología en el campo cualitativo,[9] sin ser una camisa de fuerza. Cada estudiante, tutor o investigador podrá adoptarlo o no de acuerdo con las circunstancias y naturaleza de su estudio en particular.

En el análisis de los datos, la acción esencial consiste en que recibimos datos no estructurados, a los cuales nosotros les proporcionamos una estructura. Los datos son muy variados, pero en esencia consisten en observaciones del investigador y narraciones de los participantes:[10] *a*) visuales (fotografías, videos, pinturas, entre otras), *b*) auditivas (grabaciones), *c*) textos escritos (documentos, cartas, etc.) y *d*) expresiones verbales y no verbales (como respuestas orales y gestos en una entrevista o grupo de enfoque), además de las narraciones del investigador (anotaciones o grabaciones en la bitácora de campo, ya sea una libreta o un dispositivo electrónico).

Los propósitos centrales del análisis cualitativo son: 1) explorar los datos, 2) imponerles una estructura (organizándolos en unidades y categorías), 3) describir las experiencias de los participantes según su óptica, lenguaje y expresiones; 4) descubrir los conceptos, categorías, temas y patrones presentes en los datos, así como sus vínculos, a fin de otorgarles sentido, interpretarlos y explicarlos en función del planteamiento del problema; 5) comprender en profundidad el contexto que rodea a los datos, 6) reconstruir hechos e historias, 7) vincular los resultados con el conocimiento disponible y 8) generar una teoría fundamentada en los datos.[11]

El logro de tales propósitos es una labor paulatina. Asimismo, algunas características que definen la naturaleza del análisis cualitativo son las siguientes:

1. El análisis es un proceso ecléctico (que concilia diversas perspectivas) y sistemático, mas no rígido.
2. Una fuente de datos importantísima que se agrega al análisis la constituyen las impresiones, percepciones, sentimientos y experiencias del investigador o investigadores.

[9] Por ejemplo, Savin-Baden y Major (2013), Charmaz (2013), Creswell (2013a y 2013b), Green y Thorogood (2013), Maxwell (2013), Morse (2012), Johnson y Christensen (2012), Saldaña y Leavy (2011), Denzin y Lincoln (2011), Stake (2010), Mertens (2010), The SAGE Glossary of the Social and Behavioral Sciences (2009c), Merriam (2009), Lapadat (2009), Ayres (2008), Van den Hoonaard y Van den Hoonaard (2008), Corbin y Strauss (2007), Esterberg (2002); y Strauss y Corbin (1990). Además el proceso se encuentra presente en múltiples artículos publicados en revistas como: *Qualitative Health Research*, *Qualitative Inquiry*, *Qualitative Research Journal*, *Qualitative Research in Psychology* y *Qualitative Sociology*.

[10] Morgan (2013), Lichtman (2013), Creswell (2013a), Firmin (2008), Garnham (2008) y Encyclopedia of Evaluation (2004).

[11] Creswell (2013a), Daymon (2010), Henderson (2009), Michael-Boje (2009), Baptiste (2001), Charmaz (2000); y Miles y Huberman (1994a).

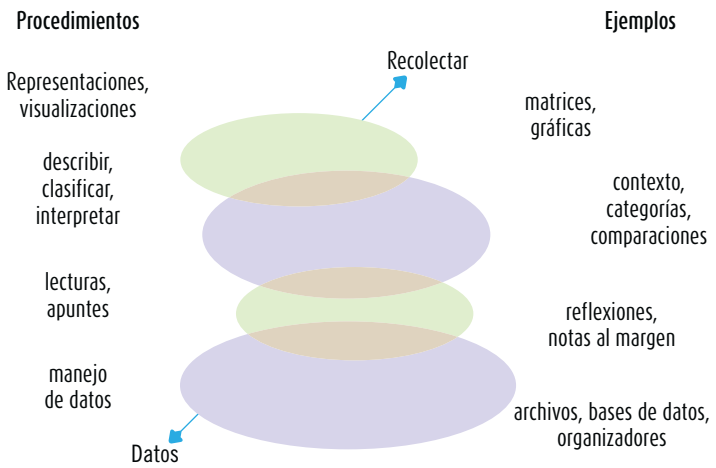
www.elosopanda.com | jamespoetrodriguez.com

3. La interpretación que se haga de los datos puede diferir de la que podrían realizar otros investigadores, lo cual no significa que una interpretación sea mejor que otra, sino que cada quien posee su propia perspectiva. Esto, pese a que recientemente se han establecido ciertos acuerdos para sistematizar en mayor medida el análisis cualitativo (Lapadat, 2009).
4. Como cualquier tipo de análisis, el cualitativo es sumamente contextual y no es un análisis "paso a paso", sino que consiste en estudiar cada dato en sí mismo y en relación con los demás ("como armar un rompecabezas").
5. Es un camino con rumbo, pero no en "línea recta", pues continuamente nos movemos; vamos y regresamos entre los primeros datos recolectados y los últimos, los interpretamos y les encontramos un significado, lo cual permite ampliar la base de datos conforme es necesario, hasta que construimos significados para el conjunto de los datos.
6. La interacción entre la recolección y el análisis nos permite mayor flexibilidad en la valoración de los datos y adaptabilidad cuando elaboramos las conclusiones (Coleman y Unrau, 2005). Debe insistirse: el análisis de los datos no es predeterminado, sino que es "prefigurado o coreografiado". Se comienza a efectuar según un plan general, pero su desarrollo sufre modificaciones de acuerdo con los resultados (Dey, 1993). El análisis es moldeado por los datos (lo que los participantes o casos van revelando y lo que el investigador va descubriendo).
7. El investigador analiza cada dato, que por sí mismo tiene un valor (Abma, 2004), y deduce similitudes y diferencias con otros datos (Rihoux, 2006).
8. Los segmentos de datos o unidades son organizados en un sistema de categorías (Ritchie, Lewis, Nicholls y Ormston, 2013; y Li y Seale, 2007).
9. Los resultados del análisis son síntesis de "orden superior" que emergen en la forma de descripciones, expresiones, categorías, temas, patrones, hipótesis y teoría (Boeije, 2009).
10. Existen diversos acercamientos al análisis cualitativo de acuerdo con el diseño o el marco referencial seleccionado. Entre estos acercamientos se encuentran varios, como etnografía, teoría fundamentada, fenomenología, feminismo, análisis del discurso, análisis conversacional, análisis semióticos y posestructurales (Grbich, 2007, y Álvarez-Gayou, 2003). Pero todos efectúan análisis temático.

Cuando después de analizar múltiples casos ya no encontramos información novedosa ("saturación"), el análisis concluye. En cambio, si se descubren inconsistencias o falta claridad en el entendimiento del problema planteado, se regresa al campo o contexto para recolectar más datos.

Creswell (1998) simboliza el desarrollo del análisis cualitativo como una espiral, en la cual se cubren varias facetas o diversos ángulos del mismo fenómeno de estudio. Esto se muestra en la figura 14.5.

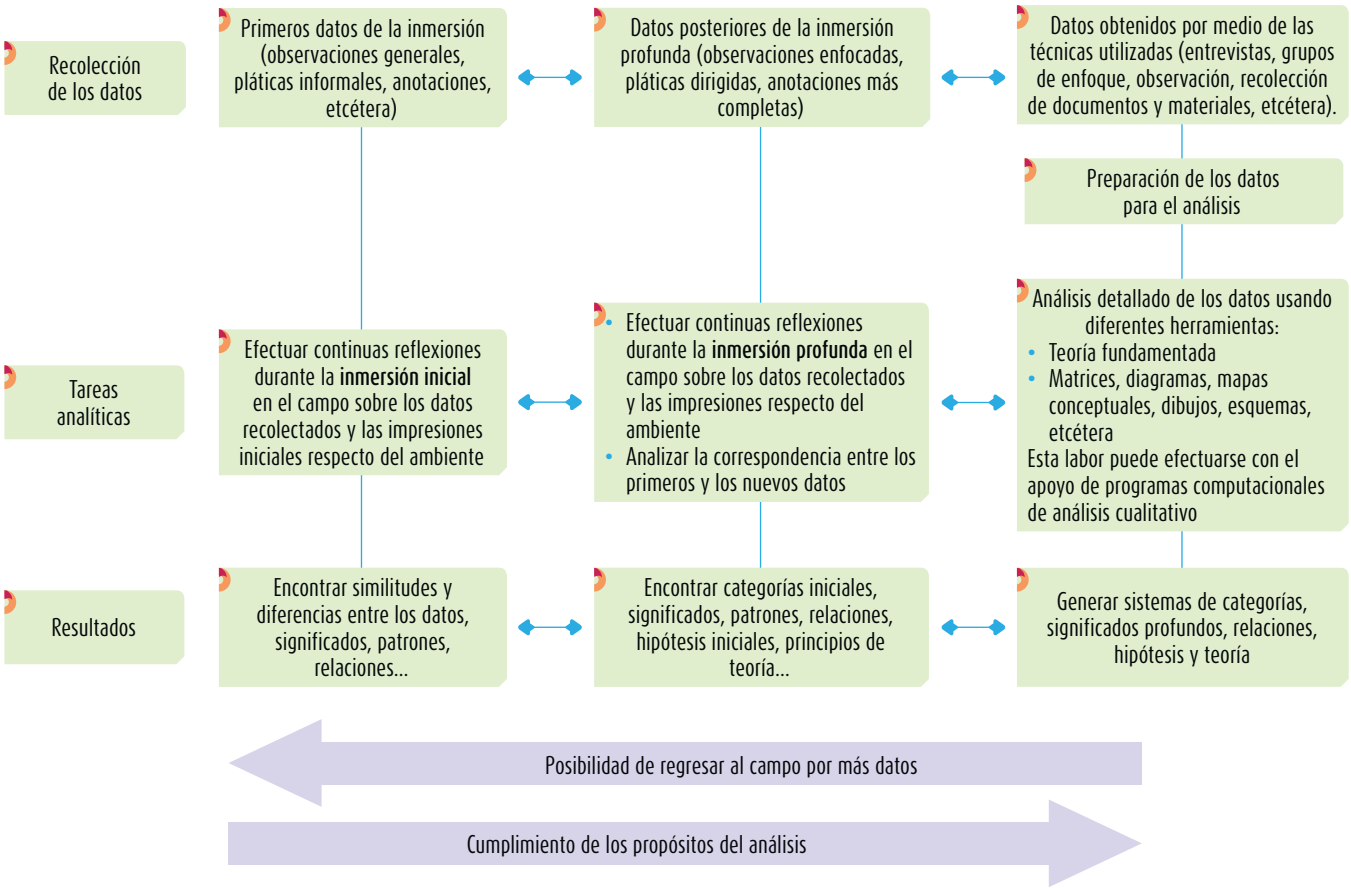
Figura 14.5 Espiral de análisis de los datos cualitativos.



Más adelante nos detendremos en algunos procedimientos de análisis (por ejemplo, las matrices).

En la figura 14.6 se presenta una propuesta de directriz general para la "coreografía" del análisis. Recordemos que se incluyen tres acciones progresivas (vertientes) en paralelo íntimamente ligadas: recolección de los datos, tareas analíticas y resultados. Las flechas en dos sentidos implican que podemos regresar a etapas previas y las flechas sin direccionalidad indican actividades asociadas.

Figura 14.6 Propuesta de "coreografía" del análisis cualitativo (directrices de las tareas potenciales para el investigador).



Ya hemos hablado de la recolección de los datos. Veamos ahora las tareas analíticas y los resultados.

Reflexiones e impresiones durante la inmersión inicial

Durante la inmersión el investigador realiza diversas observaciones del ambiente, las cuales anota junto con sus impresiones en la bitácora de campo (notas de diversos tipos, como vimos en el capítulo 12). Asimismo, el investigador platica con integrantes del ambiente (algunos de ellos son los potenciales participantes), recaba documentos y otros materiales y, en fin, realiza diversas actividades para comenzar a responder al planteamiento de su problema de investigación. Con base en estos primeros datos, el investigador reflexiona y evalúa todos los días su planteamiento (se hace preguntas como: ¿es lo que tengo en mente?, ¿el planteamiento refleja el fenómeno que quiero estudiar?, ¿el planteamiento es adecuado?, ¿debo mantenerlo o modificarlo?) y lo ajusta de acuerdo con sus propias consideraciones. También, analiza si el ambiente y la muestra son pertinentes en relación con su planteamiento y hace los cambios que crea necesarios. Como producto de las reflexiones empieza a esbozar conceptos claves que ayuden a responder al planteamiento y entender los datos (qué temas surgen, qué se relaciona con qué, qué es importante, qué se parece a qué, etcétera).

www.elosopanda.com | jamespoetrodriguez.com

Tal fue el caso del estudio de la guerra cristera. Con las primeras visitas a los contextos, los investigadores comenzaron a tener una idea de cómo fue el conflicto en cada población. Así, empezaron a entender la arraigada religiosidad de los habitantes de la región en aquellos tiempos (1926-1929) y cómo perdura hasta nuestros días.

Recordemos el ejemplo del capítulo anterior sobre los templos: en una visita a una iglesia, se encuentran agujeros en las paredes (aparentemente algo trivial), pero al visitar otro templo, el hecho se repite. Se indaga, surgen ideas (¿a qué altura están las marcas?; atraviesan de una columna a otra en los patios de los templos, ¿esto qué implica?): "aquí ataban a sus caballos, esta iglesia fue un cuartel". Se encuentra algo en común entre dos, tres o más iglesias (unidades de análisis) y al recurrir a otras fuentes, se llega a la conclusión de que los templos fueron cuarteles, en algunos casos de cristeros, pero en otros de tropas del Gobierno federal que los habían clausurado y tomado.

Reflexiones e impresiones durante la inmersión profunda

Este proceso reflexivo se mantiene conforme se recolectan más datos (¿qué me dice esto?, ¿qué significa esto otro?, ¿por qué ocurre aquello?). Las observaciones se van enfocando para responder al planteamiento, las pláticas son cada vez más dirigidas y las anotaciones más completas. En ocasiones (esto depende de la investigación en particular), se hacen las primeras entrevistas, observaciones con una guía, sesiones de grupos o recolección de materiales y objetos. Se reevalúa el planteamiento del problema, ambiente y muestra (unidades o casos). Se comparan nuevos datos con los primeros (¿en qué son similares y en qué diferentes? ¿Cómo se vinculan? ¿Qué conceptos clave se consolidan? ¿Qué otros nuevos conceptos aparecen?) De manera inductiva y paulatina surgen categorías iniciales, significados, patrones, relaciones, hipótesis primarias y principios de teoría.

Para comprender cómo avanza el análisis cualitativo (y que es casi paralelo a la recolección de los datos), tomaremos un ejemplo coloquial.

Ejemplo

Cuando vamos a conocer a una persona que nos resulta atractiva (fenómeno de estudio) en una cita en un lugar que es desconocido para nosotros, pero que fue escogido por ella o él (ambiente, contexto o escenario), ¿qué es lo primero que hacemos? Probablemente, averiguar algo de esa persona (tal vez platicamos con algunas amigas o amigos que la conocen y revisamos sus páginas en redes sociales en internet, lo que equivaldría a una revisión de la literatura). Además, buscamos información sobre el sitio. O bien, nos aventuramos y nos presentamos. Al llegar, miraremos cómo es tal lugar —inmersión inicial— (si es grande o pequeño, si tiene lugares para estacionar o parquear, cómo está decorado, si se trata de un restaurante, un bar u otro tipo, el ambiente social, etc.) y nos cuestionaremos por qué lo eligió.

Al momento de estar frente a la otra persona, la observaremos en su totalidad (desde el pelo hasta los zapatos; observación general). Comenzaremos a hacerle preguntas frecuentes (nombre, ocupación, lugar de residencia, gustos y aficiones). Mientras la observamos y conversamos (recolección de los primeros datos), nos preguntamos para nuestros adentros cómo es, qué impresión me genera, por qué me dice esto y aquello (reflexiones iniciales). Conforme transcurre la cita, iremos centrando nuestra atención en su ropa, los accesorios que trae puestos, el color de sus ojos, sus gestos [cómo sonríe, por ejemplo (observación enfocada)]; y cada vez nuestras preguntas serán más dirigidas (seguimos recolectando datos visuales y verbales, simultáneamente analizamos cada dato de manera individual y en conjunto). Al observar los movimientos de sus manos (dato), analizamos si se encuentra nerviosa o relajada (categoría) y si se trata o no de una persona expresiva (categoría). Por otro lado, establecemos relaciones entre conceptos (por ejemplo, cómo se vincula su forma de vestir con las ideas que transmite o la manera como se asocian la comunicación verbal y no verbal). Y empezamos a generar hipótesis (que emergen de los datos y la interacción misma): "es una persona calmada", "creo que podríamos ser muy buenos amigos". Y algo muy importante: fundamentamos el proceso en lo que vemos y escuchamos más que en lo que sus amigos o amigas nos dijeron de él o ella. Finalmente, hacemos preguntas más concretas y derivamos nuestras propias conclusiones. Siempre que obtenemos un dato, lo analizamos en el contexto de todo el encuentro.

Al concluir la cita, cada quien se lleva una impresión de la otra persona; ambos tienen una interpretación que es única (si la cita hubiera sido con una persona distinta, el resultado sería muy diferente). Si el encuentro fue en un restaurante, y en lugar de haber sido así, hubiera ocurrido en un bar (otro contexto), a lo mejor la situación también resultaría distinta. Así es la recolección y análisis cualitativos.

Análisis detallado de los datos

Obtuvimos los datos mediante diversas fuentes y hemos realizado reflexiones y análisis elementales. Tenemos un primer sentido de entendimiento y seguimos generando datos (cuya recolección es flexible, pero enfocada). La mayoría de las veces contamos con grandes volúmenes de datos (páginas de anotaciones u otros documentos, horas de grabación o filmación de entrevistas, sesiones grupales u observación, imágenes y distintos artefactos). ¿Qué hacer con estos datos? Como ya vimos, la forma específica de analizarlos varía según el diseño de investigación seleccionado: teoría fundamentada, estudio de caso, etnográfico, fenomenológico, etc. Cada método marca unos lineamientos para el proceso de análisis, ya que los resultados que se buscan son distintos (Grbich, 2007). Sin embargo, hay un análisis común en todo estudio cualitativo: generar categorías o temas.[12]

Teoría fundamentada Teoría o hallazgos que surgen a partir de los datos.

Para lo anterior, el procedimiento habitual es el que se ilustra en la figura 14.7 y que parte de la **teoría fundamentada** (*grounded theory*), lo cual significa que la teoría (hallazgos) va emergiendo fundamentada en los datos. Se trata de un proceso no lineal (aunque había que representarlo de alguna manera para su comprensión). Resulta sumamente iterativo (vamos y regresamos) y en ocasiones es necesario retornar al campo por más datos enfocados (entrevistas, documentos, sesiones, etcétera).

Organización de los datos y la información, así como revisión del material y preparación de los datos para el análisis detallado

Dado el amplio volumen de datos, éstos deben encontrarse muy bien organizados. Asimismo, debemos planear qué herramientas vamos a utilizar (hoy en día la gran mayoría de los análisis se efectúa mediante la computadora, al menos en un procesador de textos). Lo anterior depende del tipo de datos que hayamos generado. Pudiera ser que solamente tuviéramos datos escritos, por ejemplo, anotaciones a mano y documentos. En este caso, podemos transferir todo al procesador de textos. Si contamos únicamente con imágenes y anotaciones escritas, las primeras se escanean o transmiten a la computadora y las segundas se copian o escanean.

Cuando tenemos grabaciones de audio o video producto de entrevistas y sesiones, debemos transcribirlas para hacer un análisis exhaustivo del lenguaje (aunque algunos investigadores pueden decidir analizar directamente los materiales). La mayoría de los autores recomiendan considerar tanto las transcripciones como el análisis de primera mano. Todo depende de los recursos de que dispongamos y del equipo de investigadores con el que contemos.

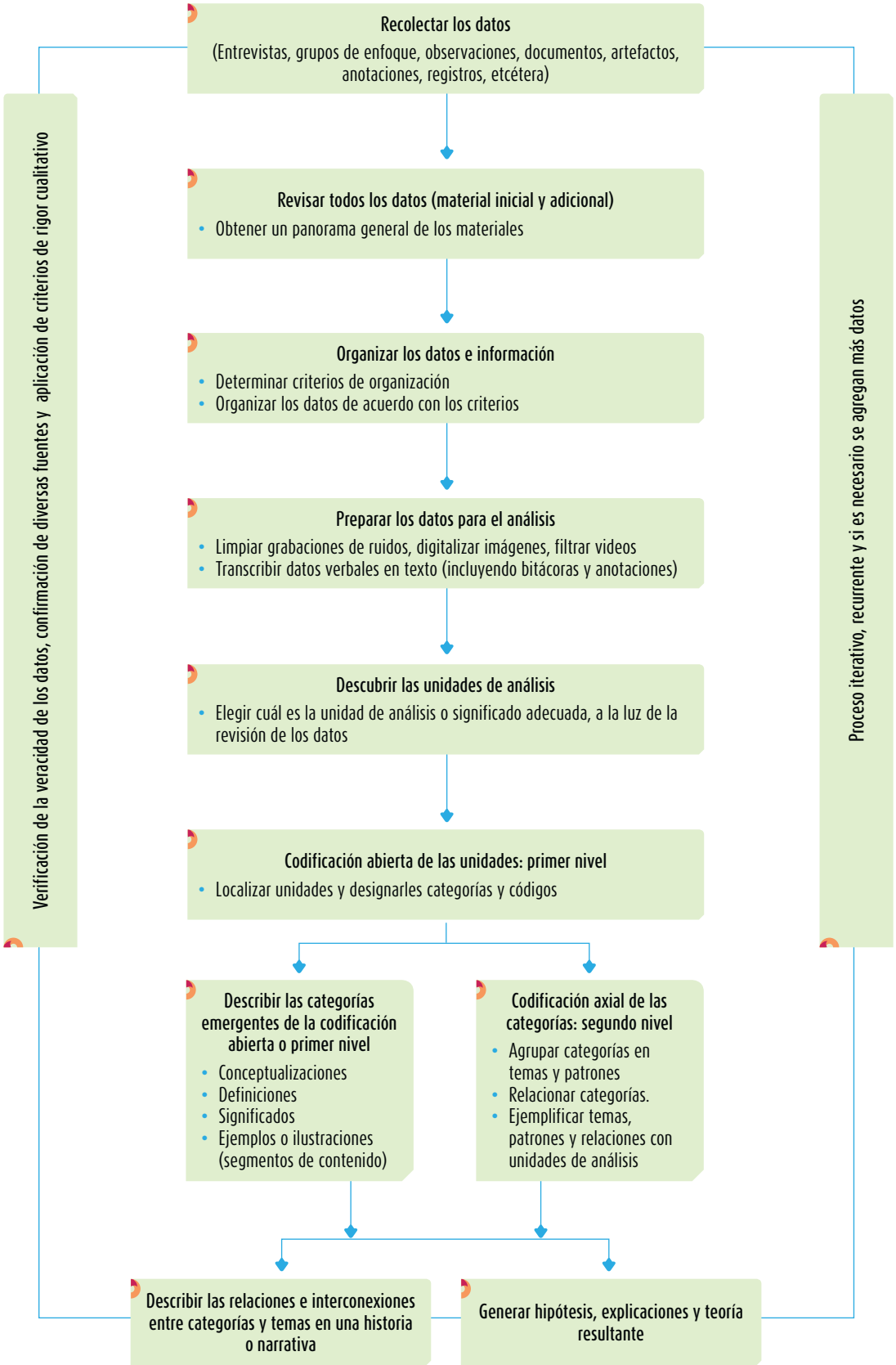
Con el material transferido a la computadora es necesario volver a examinar todo el material (explorar el sentido general de los datos) en su *forma original* (notas escritas, grabaciones en audio, fotografías, documentos, etc.). En esta revisión comenzamos a escribir una segunda bitácora (distinta a la de campo), la *bitácora de análisis*, cuya función es documentar paso a paso el proceso analítico (más adelante veremos que es una herramienta fundamental). Durante tal escrutinio debemos asegurar que el material esté completo y posea la calidad necesaria para ser analizado; en caso de que no sea así (grabaciones que no se entienden, documentos que no pueden leerse), es preciso realizar las mejoras técnicas posibles (depurar las grabaciones, optimizar imágenes, etcétera).

La tarea de transcripción de los materiales grabados es compleja y requiere paciencia. Por ejemplo, una hora de entrevista da aproximadamente 30 a 50 páginas en el procesador de textos (depende del programa, márgenes e interlineado). Y lleva más o menos de tres a cuatro horas. Si se dispone de varias personas para esta labor, el investigador puede realizar dos o tres transcripciones para mostrar reglas y procedimientos (Coleman y Unrau, 2005). Quienes transcriban deberán capacitarse (el número de personas depende del volumen de datos, los recursos y el tiempo que tengamos para completar las transcripciones). A continuación, revisaremos una serie de recomendaciones para la hora de transcribir.

[12] Creswell (2013a); Savin-Baden y Major (2013); Melanie-Birks y Mills (2011); Floersch, Longhofer, Kranke y Townsend (2010); Mertens (2010); Lapadat (2009); O'Neil (2008); Noerager (2008); Charmaz (2008); Wiersma y Jurs (2008); Torrance (2001); Denzin y Lincoln (2000); y Strauss y Corbin (1998).

www.elosopanda.com | jamespoetrodriguez.com

Figura 14.7 Proceso de análisis cualitativo para generar categorías o temas.



- Por cuestiones de ética, es preciso apegarse al principio de confidencialidad. Para esto, es posible sustituir el nombre verdadero de los participantes por códigos, números, iniciales, apodos u otros nombres, tal como hicieron Morrow y Smith (1995). Lo mismo ocurre con el reporte de resultados.
- Utilizar un formato con márgenes amplios y separar las intervenciones con doble espacio (por si queremos hacer anotaciones o comentarios). Por ejemplo: en entrevistas, las intervenciones del entrevistador y del entrevistado; en sesiones, la intervención del conductor y de cada participante (cada vez que alguien interviene), señalando quién realiza la participación:

  Entrevistador: ¿Me podrías aclarar el punto?
  Entrevistado: Desde luego que sí. Ana Paola siempre me ha parecido atractiva; si no le he propuesto ir más allá es porque…
  Entrevistador: Pero, entonces, ¿cómo podrías definir tu relación con ella?
  Entrevistado: es algo diferente, extraña, dadas las circunstancias…

  Es decir, indicar cuándo comienza y termina cada pregunta y respuesta.
- Transcribir todas las palabras, sonidos y elementos paralingüísticos: muecas, interjecciones (como ¡oh!, ¡mmm!, ¡eh! y demás).[13]
- Indicar pausas (pausa) o silencios (silencio); expresiones significativas (llanto, risas, golpe en la mesa); sonidos ambientales (timbró el teléfono móvil; se azotó la puerta); hechos que se deduzcan (entró alguien); cuando no se escucha (inaudible), etc. Se trata de incluir el máximo de información.
- Si vamos a analizar línea por línea (cuando ésta va a ser la unidad de análisis), numerar todos los renglones (lo que se puede hacer automáticamente en los procesadores de texto y los programas de análisis cualitativo).

Una vez transcritos los materiales mediante el debido equipo, lo ideal es volver a revisar todos los datos, ahora *reprocesados* (incluso anotaciones), en particular, si varios investigadores los recolectaron. De cualquier manera, ayuda a recordar casos y vivencias en el campo (Coleman y Unrau, 2005). En este momento, leemos y releemos varias veces las transcripciones para profundizar el sentido general de los datos, al mismo tiempo que comenzamos a cuestionarnos: ¿qué ideas generales mencionan los participantes? ¿Qué tono tienen dichas ideas? ¿Qué me dicen los datos? (Creswell, 2013a).

La siguiente actividad es organizar con precisión los datos, mediante algún criterio o varios que creamos más convenientes.

Algunos de estos criterios son:

1. Cronológico (por ejemplo, orden en que fueron recolectados: por día y bloque, como mañana, tarde y noche).
2. Por sucesión de eventos (por ejemplo, en el caso de una catástrofe, como una inundación o terremoto: antes de la calamidad, durante ella e inmediatamente después de la catástrofe —digamos hasta que cesaron los efectos físicos— y etapa posterior (secuela).
3. Por tipo de datos: entrevistas, observaciones, documentos, fotografías, artefactos, etcétera.
4. Por grupo o participante (por ejemplo: Marcela, Lucy, Ana Paola, Roberto, Sergio…; mujeres y hombres; médicos, enfermeras, paramédicos, pacientes, familiares…).
5. Por ubicación del ambiente (centro de la catástrofe, cercanía, periferia, lejanía).
6. Por tema (por ejemplo, en un estudio sobre las relaciones en un hospital, si hubo sesiones en los que la discusión se centró en el tema de la seguridad en el recinto, mientras que en otras se trató sobre la calidad en la atención y en algunas más sobre los problemas emocionales de los pacientes).
7. Importancia del participante (testimonios de actores principales, testimonios de actores secundarios).

   O bien otro criterio.

[13] En cada país y región hay expresiones propias de la cultura local.

www.elosopanda.com | jamespoetrodriguez.com

A veces los datos también se organizan mediante varios criterios progresivos; por ejemplo, primero por tipo (**transcripciones** de entrevistas y anotaciones), y luego estas últimas por la clase de notas (de la observación, interpretativas, temáticas, personales y de reactividad); o criterios cruzados (combinaciones). Un ejemplo lo sería la matriz que se presenta en la tabla 14.8.

Transcripción Registro escrito de una entrevista, sesión grupal, narración, anotación y otros elementos similares. Refleja el lenguaje verbal, no verbal y contextual de los datos.

Tabla 14.8 Guerra cristera

	Municipio	Irapuato	Cortazar	Villagrán	Salamanca	Apaseo
Materiales	Notas Observaciones Documentos Entrevistas Fotografías					

En el caso de documentos, materiales, artefactos, grabaciones, etc., es conveniente elaborar un listado o relación que los contenga a todos (con número, fecha de realización, fecha de transcripción y aquellos otros datos apropiados). Además, no debemos olvidar respaldar todo documento en al menos dos fuentes (tener una copia adicional de las fotografías y grabaciones).

Bitácora de análisis

Esta bitácora tiene la función de documentar el procedimiento de análisis y las reacciones del investigador al proceso y contiene fundamentalmente:

- Anotaciones sobre el método utilizado (se describe el proceso y todas las actividades realizadas; por ejemplo: ajustes a la codificación, problemas y la forma en que se resolvieron).
- Anotaciones sobre ideas, conceptos, significados, categorías e hipótesis que van surgiendo.
- Anotaciones en relación con la credibilidad y verificación del estudio, para que cualquier otro investigador pueda evaluar el trabajo (información contradictoria, razones por las cuales se procede de una u otra forma).

Resulta ser un instrumento invaluable para la validez y confiabilidad del análisis (punto que abordaremos al final del capítulo).

Cuando realizamos la codificación o categorización de los datos, pueden surgir interrogantes, ideas, hipótesis y conceptos que nos comiencen a ilustrar en torno al planteamiento del problema, por lo que resulta indispensable escribirlos para que no olvidemos cuestiones importantes. Las notas nos ayudan a identificar unidades y categorías de significado. Es una estrategia útil para organizar los procedimientos analíticos. Para las anotaciones, que suelen también llamarse "memos analíticos", Strauss y Corbin (1998) recomiendan:

- Registrar la fecha de la anotación o memorándum.
- Incluir cualquier referencia o fuente importante (por ejemplo, si consultamos con un colega, a quién, de qué institución y cuál es su comentario).
- Marcar los memos con encabezados que sinteticen la idea, categoría o concepto señalado.
- No restringir el contenido de los memos o anotaciones; permitirnos el libre flujo de ideas.
- Identificar el código al cual pertenece el memo.
- Usar diagramas, esquemas y matrices (u otra clase de síntesis analítica) en los memos para explicar ideas, hipótesis y conceptos.
- Cuando uno piense que una categoría o un concepto quedó lo suficientemente definido, crear un memo adicional (distinguirlo y etiquetarlo con la palabra "saturación").
- Registrar las reflexiones en memos que ayuden a pasar de un nivel descriptivo a otro interpretativo.
- Guardar una copia de todos los memos.

Memo analítico Documenta decisiones o definiciones hechas al momento de analizar los datos. Desde cómo surge una categoría hasta el código que se le asigna o el establecimiento de una regla de codificación.

Los **memos analíticos** se elaboran con fines de triangulación o auditoría entre investigadores, para que otras personas puedan ver lo que hicimos y cómo lo hicimos (Coleman y Unrau, 2005).

La bitácora se escribe diariamente (anotando la fecha) y cada investigador sigue su propio sistema para llenarla. Grinnell y Unrau (2007) proponen el siguiente esquema: 1) memos, anotaciones o comentarios acerca del método de análisis, 2) memos sobre los problemas durante el proceso, 3) memos en relación con la codificación, 4) memos respecto a ideas y comentarios de los investigadores (incluyendo diagramas, mapas conceptuales, dibujos, matrices, etc., 5) memos sobre el material de apoyo localizado (fotografías, videos, etc.) y 6) memos relacionados con significados, descripciones y conclusiones preliminares.

Así como la bitácora de campo refleja lo que "transpiramos" durante la recolección de los datos y nos ayuda a establecer la credibilidad de los participantes, la bitácora analítica muestra lo que "transpiramos" al analizar los datos y nos apoya a establecer la credibilidad del método de análisis.

Surgimiento de unidades de análisis y codificación abierta (en primer nivel o plano inicial)

En la mayoría de los estudios cualitativos se codifican los datos para tener una descripción más completa de éstos, se resumen, se elimina la información irrelevante y se realizan análisis cuantitativos elementales; finalmente, se trata de entender mejor el material analizado.

La codificación tiene dos planos o niveles: en el primero (codificación abierta), se codifican las unidades (datos en bruto) en categorías; en el segundo, se comparan las categorías entre sí para agruparlas en temas y buscar posibles vinculaciones (Saldaña, 2012; Matthew y Price, 2009a; Wicks, 2009; y Miles y Huberman, 1994).

El primer nivel es una combinación de varias acciones: identificar unidades de significado, categorizarlas y asignarles códigos a las categorías. A diferencia de la codificación cuantitativa, en la que una unidad constante se ubica en un sistema de categorías, en la codificación cualitativa abierta el investigador considera un segmento o unidad de contenido (no siempre estándar), lo analiza (se cuestiona: ¿qué significa este segmento?, ¿a qué se refiere?, ¿qué me dice?); toma otro segmento, también lo analiza, compara ambos segmentos y los analiza en términos de similitudes y diferencias (¿qué significado tiene cada uno?, ¿qué tienen en común?, ¿en qué difieren?, ¿me dicen lo mismo o no?). Si los segmentos son distintos en términos de significado y concepto, de cada uno induce una categoría (o bien, considera que no poseen un significado para el planteamiento); si son similares, induce una categoría común. Al considerar un tercer segmento, el investigador lo analiza conceptualmente y en términos de significado; del mismo modo, lo contrasta con los dos anteriores, evalúa similitudes y diferencias, induce una nueva categoría o lo agrupa con los otros. Con el cuarto segmento repite el proceso, y así sucesivamente (a este procedimiento se le denomina "comparación constante") (Tracy, 2013; Noerager, 2008; y Parry, 2003). El investigador va otorgando significados a los segmentos o unidades y descubriendo categorías. A cada una de éstas les asigna un código.

Codificación cualitativa El investigador considera segmentos de contenido, los analiza y compara. Si son distintos en términos de significado y concepto, de cada uno induce una categoría; si son similares, induce una categoría común.

Categorías Conceptualizaciones analíticas desarrolladas por el investigador para organizar los resultados o descubrimientos relacionados con un fenómeno o experiencia humana que está bajo investigación.

Códigos Identifican las categorías que emergen de la comparación constante de segmentos o unidades de análisis.

En la **codificación cualitativa** los códigos surgen de los datos (más precisamente, de los segmentos de datos): los datos van mostrándose y los "capturamos" en categorías. Usamos la codificación para comenzar a revelar significados potenciales y desarrollar ideas, conceptos e hipótesis; vamos comprendiendo lo que sucede con los datos (empezamos a generar un sentido de entendimiento respecto al planteamiento del problema). Los **códigos** son etiquetas para identificar **categorías**, es decir, describen un segmento de texto, imagen, artefacto u otro material.

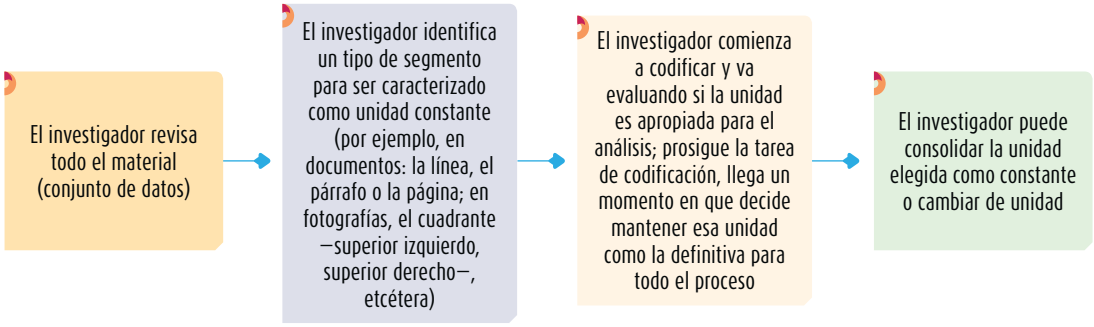
Cuando consideramos que un segmento o unidad es relevante (en términos del planteamiento, de representatividad de lo que expresaron los participantes, de importancia a juicio del investigador) podemos extraerlo como un potencial ejemplo de la categoría o de los datos.

Conforme el investigador revisa nuevos segmentos de datos y vuelve a revisar los anteriores segmentos (comparación constante), continúa "conectando conceptualmente" unidades y genera más categorías o consolida las anteriores.

www.elosopanda.com | jamespoetrodriguez.com

Cabe señalar que la identificación de unidades o segmentos es tentativa en su comienzo y se encuentra sujeta a cambios. En la literatura sobre investigación cualitativa podemos identificar dos maneras para definir las unidades de análisis que serán codificadas: la primera, que podemos denominar como la elección de una "unidad constante", implica el proceso que se muestra en la figura 14.8.

Figura 14.8 Proceso de elección de una unidad constante.



La segunda, la de "libre flujo" implica que las unidades no poseen un tamaño equivalente. Se selecciona el inicio del segmento y hasta que se encuentra un significado, se determina el final del segmento. Por ejemplo, algunos segmentos podrían tener cinco líneas, otros 10, otros 50 (desde luego, en el caso del párrafo o una intervención de un participante en una sesión de grupo, como unidad constante, puede poseer diferentes extensiones también).

En ambas formas para decidir cuál es la unidad o segmento de análisis, es posible cambiar de unidad en cualquier momento. Incluso, quizá decidamos utilizar en un mismo estudio las dos posibilidades para diferentes clases de datos (entrevistas y fotografías, por ejemplo).

Codificación Implica, además de identificar experiencias o conceptos en segmentos de los datos (unidades), tomar decisiones acerca de qué piezas "embonan" entre sí para ser categorizadas, codificadas, clasificadas y agrupadas para conformar los patrones que serán empleados con el fin de interpretar los datos.

En este primer nivel de análisis, las categorías (y códigos) identificadas deben relacionarse lógicamente con los datos que representan (que quede clara la vinculación). Las categorías pueden emerger de preguntas y reflexiones del investigador o reflejar los eventos críticos de las narraciones de los participantes.

En la bitácora de análisis es necesario explicar con claridad las razones por las que se genera una categoría.

La esencia del proceso reside en que a segmentos que comparten naturaleza, significado y características, se les asigna la misma categoría y código, los que son distintos se ubican en diferentes categorías y se les proporcionan otros códigos. La tarea es identificar y etiquetar categorías relevantes de los datos.

Los segmentos se convierten en unidades cuando poseen un significado (de acuerdo con el planteamiento del problema) y en categorías del esquema final de codificación en el primer nivel, si su esencia se repite más adelante en los datos (por ejemplo, en la entrevista o en otras entrevistas). Las unidades son segmentos de los datos que constituyen los "tabiques" para construir el esquema de clasificación y el investigador considera que tienen un significado por sí mismas.

Bazeley (2013) y Coffey y Atkinson (1996) señalan que son tres las actividades de la codificación abierta o en primer plano:

1. Advertir cuestiones relevantes en los datos.
2. Analizar esas cuestiones para descubrir similitudes y diferencias, así como estructuras.
3. Recuperar ejemplos de tales cuestiones (unidades o segmentos).

Tal como resumen Coleman y Unrau (2005), la codificación en el primer nivel es predominantemente concreta y consiste en identificar propiedades de los datos. Las categorías se forman comparando datos, pero en este nivel no combinamos ni relacionamos datos. Todavía no interpretamos el significado subyacente en ellos. En la codificación abierta se trata intensivamente, unidad por unidad,

con la identificación de categorías que pudieran ser interesantes, sin limitarnos. Es importante asegurarnos de entender las categorías que se muestran en los datos.

En el caso de que el investigador decida la elección de una "unidad constante", algunos ejemplos serían los que a continuación se mencionan.

En textos (documentos, notas, transcripciones, etcétera):

1. Palabras: "alcoholismo", "Ricardo", "divorcio", "patria".
2. Líneas: "mi esposo me abandonó después de que me embaracé por tercera vez"
3. Párrafos:

   No puedo dejar de pensar (mmm) en que mis hijas vean a su padre completamente ebrio. Es algo en lo que pienso todas las noches antes de acostarme. Ojalá dejara la bebida (¡uhh!), pero lo veo como algo imposible. No ha podido dejar de hacerlo desde que lo conozco… (¡mmm!) pero antes bebía mucho menos (…)

4. Intervenciones de participantes (desde que comienza hasta que concluye su intervención cada uno).

   Jesús: No puedo dejar de beber, no puedo (¡ugg!).
   Alejandra: Ni lo quieres intentar. Piensa en lo mal que te sientes; en tus hijas. ¿Qué va a pasar cuando sean grandes?

   En este último caso tenemos dos unidades de análisis (intervenciones).

5. Páginas.
6. Cambios de tema (cada vez que aparece un nuevo tema).
7. Todo el texto.

En grabaciones de audio o video (hayan o no sido transcritas a texto):

1) Palabras o expresiones, 2) intervenciones de participantes, 3) cambios de tema, 4) periodos (segundo, minuto, cada *k* minutos, hora), o 5) sesión completa (entrevista, grupo de enfoque, otro).

En biografías: 1) periodo (día, mes, año, etapa o pasaje de vida), 2) cambios de tema y 3) actos conocidos.

Música: 1) verso de canción, 2) estrofa, 3) canción completa y 4) obra.

Construcciones, materiales o artefactos: 1) pieza completa, 2) partes específicas o sitios dependiendo del material, artefacto o construcción (en iglesias: atrio, altar, confesionario, etcétera).

A continuación presentamos muestras de unidades de análisis de los principales ejemplos que se han ido desarrollando en esta tercera parte del libro (tabla 14.9).

Tabla 14.9 Muestras de unidades de significado en los ejemplos desarrollados

Estudio	Participantes	Método de recolección de los datos	Ejemplos de unidades
Estudio sobre las experiencias de abuso sexual infantil de Morrow y Smith (1995)	Mujeres adultas que habían experimentado abuso sexual durante su infancia	Entrevista inicial y sesiones de enfoque	• "Solía jugar con muñecas de papel. Ellas eran mis amigas. Ellas nunca me podrían lastimar". • "Yo me refugié en la abuela, ella era una mujer muy espiritual... Ella acostumbraba mecernos y cantarnos". • "Me aislé para siempre". • "Debo ser invisible siendo buena niña, muy buena niña".
La guerra cristera en Guanajuato (1926-1929)	Sobrevivientes de la época	Entrevista	• El padre poco a poco platicaba: "no estamos bien aquí, niño, de ninguna manera... Mira, aquí me agarran con los cabecillas... Pedro no tiene ni para cuándo irse... no saben que aquí está el cabecilla, porque si no...". • "Estaba duro, si llegaba a los ranchos el Gobierno, se tragaba lo que había ahí y la gente se quedaba con hambre, si llegaban los cristeros, igual, no, se armó un desmadre". • "Todo valía madres, el que les caía mal, lo mataban. Había quienes eran cristeros y no ponían ni una pata en la cárcel".

(*continúa*)

Tabla 14.9 (*continuación*)

Estudio	Participantes	Método de recolección de los datos	Ejemplos de unidades
Evaluación de la experiencia de compra de los clientes en centros comerciales de una importante cadena latinoamericana	Clientes de diferentes edades	Grupos de enfoque	• "Yo vengo a comprar, cuando tenemos tiempo nos venimos a tomar un café, pero deberían de abrir más temprano, mis hijos que son adolescentes vienen por cercanía y visitan mucho la parte del *fast-food*, en la tarde hay mucho joven". • "El ambiente me da paz, huele rico, el sonido me gusta, es agradable el aire acondicionado y todo está muy ordenado". • "La plaza no es para comprar, es más bien para dar la vuelta".

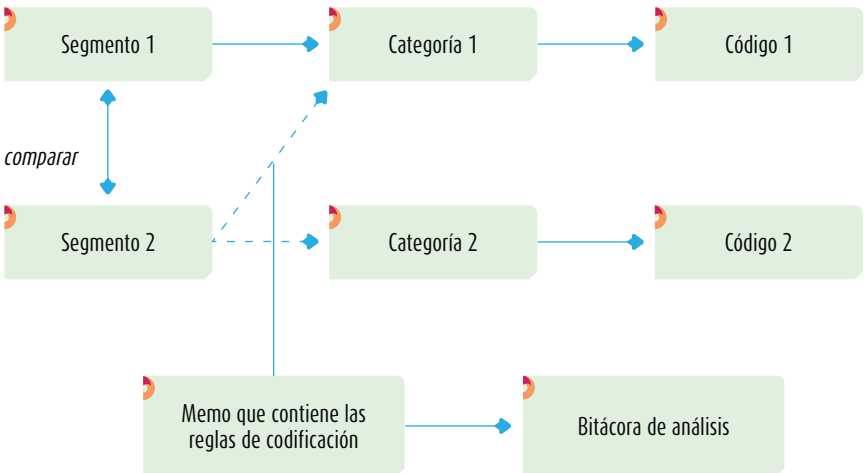
Las unidades o segmentos de significado se analizan tal como se recolectan en el campo (en el lenguaje de los participantes, aunque las expresiones sean gramaticalmente incorrectas, la estructura sea incoherente, haya faltas de ortografía e incluso groserías o términos vulgares).

En nuestra bitácora de análisis, por ejemplo, anotaríamos que las unidades de análisis elegidas fueron las líneas. Asimismo, recordemos que en la codificación cualitativa, las **categorías** son conceptos, experiencias, ideas, hechos relevantes y con significado.

Categorías Deben guardar una relación estrecha con los datos.

En resumen, desde el comienzo (a través de la comparación constante), cada segmento o unidad se clasifica como similar o diferente de otras. Si las primeras dos unidades poseen cualidades similares, generan tentativamente una categoría y a ambas se les asigna un mismo código. En el momento de asignar los códigos, elaboramos una nota sobre las características de las unidades por las que se consideran similares (un memo analítico sobre la regla), la cual se incluye en la bitácora de análisis. Si las dos unidades no son similares, la segunda produce una nueva categoría y se le asigna otro código. Y de nuevo, la información que defina a esta segunda categoría, se registra en la bitácora (regla en el memo analítico). Durante el proceso se van especificando las reglas que señalan cuándo y por qué se incluye una unidad en esa categoría. Tomamos una tercera, cuarta, quinta, *k* unidad o segmento y repetimos el proceso. La actividad se esquematiza en la figura 14.9.

Figura 14.9 Proceso de codificación cualitativa abierta.



El número de categorías aumenta cada vez que el investigador identifica unidades diferentes (en cuanto a significado) del resto de los datos (unidades previas categorizadas).

Veámoslo con un ejemplo sencillo, el análisis de los tipos de violencia entre parejas, en que se muestra una transcripción con nueve unidades de análisis (constantes), definiendo a la unidad como línea:

Ejemplo

De unidades de análisis (constantes)

Investigación sobre la experiencia negativa de una mujer golpeada por su esposo y los tipos de violencia que ejercen los maridos que abusan de sus parejas.

Recolección de los datos: Entrevistas en profundidad.
Unidad de análisis: línea.
Contexto: entrevista con una joven esposa de 20 años, dos años de casada, de origen humilde, que vive en los suburbios de Valledupar, Colombia.

1. Carolina: Mi esposo me ha golpeado varias veces (¡ehh!) (pausa).
2. No sé cómo decirlo. Me pega con la mano abierta y con el puño.
3. La última vez me dijo: "Eres una ramera". También me ha
4. dicho que soy malnacida, perra. Siempre me insulta. Y la
5. verdad es que nunca he dado motivo. Nunca (pausa). Me dice que
6. los hombres se meten en mí como culebras. Que me gusta hacerlo
7. quedar mal. Me mira con odio del malo. Me amenaza con los ojos.
8. Y a veces le contesto y le pego también. El otro día le rompí
9. una lámpara en la cabeza...

Analizamos la primera línea (unidad de análisis):

1. Carolina: "Mi esposo me ha golpeado varias veces (¡ehh!) (pausa)".

Consideramos su significado: ¿a qué se refiere? Decidimos generar la categoría "violencia física" (memo: la "violencia física" implica que una persona arremete contra la otra utilizando una parte de su cuerpo). Si más adelante encontramos que en la violencia física se utilizan objetos (además de partes del cuerpo), la regla podría modificarse (ampliándola): la "violencia física" implica que una persona arremete contra la otra utilizando una parte de su cuerpo o un objeto. Y si resulta que se producen hematomas o heridas, podría agregarse a la regla, lo mismo que "uso de armas de fuego". Cada elemento nuevo se agrega a la regla o definición.

La segunda unidad o segmento:

2. "No sé cómo decirlo. Me pega con la mano abierta y con el puño".

La comparamos con la primera (¿significan ambas lo mismo?, ¿qué clase de violencia reflejan?). La conclusión es que se refiere a lo mismo, sería parte también de "violencia física".

La tercera:

3. "La última vez me dijo: 'Eres una ramera'. También me ha…"

¿Qué significa? ¿Comparada con las otras dos significa lo mismo? La respuesta es que resulta ser algo diferente, pues no se trata de violencia física y no aplica la regla establecida. Esta tercera unidad posee un significado distinto, así que creamos la categoría "violencia verbal" (memo: "la violencia verbal" se refiere a que una persona insulta a la otra).

El cuarto segmento o unidad:

4. "…dicho que soy malnacida, perra. Siempre me insulta. Y la…"

¿Qué significa? (comparada esta unidad con las demás, ¿es similar o diferente?) La respuesta es que es distinta de las primeras dos y similar a la tercera, por lo que se asigna a la categoría "violencia verbal".

La quinta:

5. "…verdad es que nunca he dado motivo. Nunca (pausa). Me dice que…"

No es similar a ninguna unidad; debe crearse otra categoría, pero si el análisis está dirigido a describir los tipos de violencia utilizados por el marido, tal unidad no es pertinente para generar categorías.

Sin embargo, si el análisis pretende evaluar, además de los tipos de violencia presentes en las interacciones, el contexto en que se dan y la atribución de la esposa respecto de las razones por las cuales

5.4.2. Representación de datos mediante tablas y gráficos

Para el desarrollo del subtema **Representación de datos mediante tablas y gráficos**, se consideran aquellos contenidos relacionados con la organización, descripción y presentación de los datos cuantitativos mediante distribuciones de frecuencias y representaciones gráficas, de manera que la información obtenida en una investigación pueda ser organizada, resumida, visualizada e interpretada. En la obra de Roberto Hernández Sampieri, estos contenidos se desarrollan principalmente en el **Capítulo 10: Análisis de datos cuantitativos**, específicamente en los apartados referidos a la estadística descriptiva, las distribuciones de frecuencias y su representación gráfica.

N.°	Libros	Capítulo	Página	Contenido
1	*Metodología de la investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**p. 270**	Análisis de datos cuantitativos. **Introducción al procedimiento para analizar cuantitativamente los datos.**
2	*Metodología de la investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**p. 282**	**Estadística descriptiva para cada variable.** Organización y descripción de los datos correspondientes a cada variable.
3	*Metodología de la investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**p. 282**	**¿Qué es una distribución de frecuencias?** Organización de los datos mediante distribuciones de frecuencias.
4	*Metodología de la investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**p. 283**	**¿Qué otros elementos contiene una distribución de frecuencias?** Elementos que complementan la organización de los datos en una distribución de frecuencias.
5	*Metodología de la investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**p. 284**	**¿De qué otra manera pueden presentarse las distribuciones de frecuencias?** Formas de presentación de las distribuciones de frecuencias.
6	*Metodología de la investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**p. 284**	**Las distribuciones de frecuencias también se pueden graficar como polígonos de frecuencias.** Representación gráfica de las distribuciones de frecuencias.
7	*Metodología de la investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**p. 286**	**¿Cuáles son las medidas de tendencia central?** Medidas descriptivas utilizadas para resumir los datos.

8	*Metodología de la investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**p. 287**	**¿Cuáles son las medidas de la variabilidad?** Medidas utilizadas para describir la variabilidad de los datos.
9	*Metodología de la investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**p. 288**	**¿Cómo se interpretan las medidas de tendencia central y de la variabilidad?** Interpretación de las medidas estadísticas descriptivas.
10	*Metodología de la investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**p. 327**	**Paso 7: preparar los resultados para presentarlos.** Preparación y presentación de los resultados de la investigación.

5.4.2. Representación de datos mediante tablas y gráficos

Para el desarrollo del subtema **Representación de datos mediante tablas y gráficos**, se consideran aquellos contenidos orientados a la organización, resumen, exploración y representación gráfica de los datos estadísticos, mediante distribuciones de frecuencias, histogramas y diferentes tipos de gráficas. Asimismo, se consideran aspectos relacionados con la correcta presentación de la información gráfica y la identificación de representaciones que pueden informar adecuadamente o inducir a interpretaciones erróneas. De manera complementaria, se incluyen los diagramas de dispersión y las gráficas de caja como recursos para representar y explorar visualmente los datos.

N.°	Libros	Capítulo	Página	Contenido
1	*Estadística* — Mario F. Triola	**Capítulo 2. Exploración de datos con tablas y gráficas**	**p. 40**	Exploración de datos mediante tablas y gráficas. Introducción a la organización y representación de datos.
2	*Estadística* — Mario F. Triola	**Capítulo 2. Exploración de datos con tablas y gráficas — 2-1**	**p. 42**	**Distribuciones de frecuencias para organizar y resumir datos.** Organización y resumen de los datos mediante distribuciones de frecuencias.
3	*Estadística* — Mario F. Triola	**Capítulo 2. Exploración de datos con tablas y gráficas — 2-2**	**p. 51**	**Histogramas.** Representación gráfica de datos mediante histogramas.
4	*Estadística* — Mario F. Triola	**Capítulo 2. Exploración de datos con tablas y gráficas — 2-3**	**p. 57**	**Gráficas que informan y gráficas que engañan.** Consideraciones para la elaboración e interpretación adecuada de representaciones gráficas.
5	*Estadística* — Mario F. Triola	**Capítulo 2. Exploración de datos con tablas y gráficas — 2-4**	**p. 67**	**Diagramas de dispersión, correlación y regresión.** Representación gráfica de la relación entre variables.
6	*Estadística* — Mario F. Triola	**Capítulo 3. Descripción, exploración y comparación de datos — 3-3**	**p. 112**	**Medidas de posición relativa y gráficas de caja.** Representación gráfica mediante gráficas de caja como complemento para la exploración de datos.

2-1 Distribuciones de frecuencias para organizar y resumir datos

2-2 Histogramas

2-3 Gráficas que informan y gráficas que engañan

2-4 Diagramas de dispersión, correlación y regresión

2

EXPLORACIÓN DE DATOS CON TABLAS Y GRÁFICAS

PROBLEMA DEL CAPÍTULO

Restaurantes de comida rápida: ¿Cuál es el más rápido?

Una atractiva característica de los restaurantes de comida rápida es ¡que son rápidos! Para seguir siendo competitivos, los restaurantes de comida rápida deben no sólo proporcionar una buena experiencia culinaria, sino también deben hacerlo tan rápido como sus competidores. El conjunto de datos 25 "Comida rápida" del apéndice B muestra los tiempos de servicio en el auto (medidos en segundos), que se obtuvieron de muestras de clientes en diferentes restaurantes. En la tabla 2-1 se listan las 50 mediciones del tiempo de servicio de la primera columna del conjunto de datos 25. Quien pueda obtener conclusiones significativas simplemente con mirar esos datos es una persona excepcionalmente rara. En este capítulo presentamos métodos que se enfocan en

organizar y resumir los datos y utilizar gráficas que permitan entender las características importantes de los datos, especialmente su distribución. Estos métodos nos ayudarán a comparar los restaurantes.

TABLA 2-1 Tiempos de servicio en el auto (segundos) para los almuerzos en McDonald's

107	139	197	209	281	254	163	150	127	308	206	187	169	83	127	133	140
143	130	144	91	113	153	255	252	200	117	167	148	184	123	153	155	154
100	117	101	138	186	196	146	90	144	119	135	151	197	171	190	169	

OBJETIVOS DEL CAPÍTULO

Este capítulo y el siguiente se centran en las características más importantes de los datos, entre las que se incluyen:

Características de los datos

1. **Centro:** Un valor representativo que indica dónde se encuentra el centro del conjunto de datos.
2. **Variación:** Una medida de qué tanto varían los valores de los datos.
3. **Distribución:** La naturaleza o forma de la dispersión de los datos en el rango de valores (por ejemplo, en forma de campana).
4. **Datos atípicos:** Valores muestrales que están muy alejados de la gran mayoría de los demás valores de la muestra.
5. **Tiempo:** Cualquier cambio en las características de los datos a través del tiempo.

Este capítulo proporciona herramientas que permiten conocer los datos al organizarlos, resumirlos y representarlos de manera que sea posible observar sus características más importantes. Los objetivos del capítulo son:

2-1 Distribuciones de frecuencias para organizar y resumir datos

- Desarrollar la capacidad de resumir datos en el formato de una distribución de frecuencias y una distribución de frecuencias relativas.
- Para una distribución de frecuencias, identificar los valores de la anchura de clase, la marca de clase, los límites de clase y las fronteras de clase.

2-2 Histogramas

- Desarrollar la capacidad de representar la distribución de datos en el formato de un histograma o un histograma de frecuencias relativas.
- Examinar un histograma e identificar las distribuciones comunes, incluyendo una distribución uniforme y una distribución normal.

2-3 Gráficas que informan y gráficas que engañan

- Desarrollar la capacidad de graficar datos utilizando un diagrama de puntos, una gráfica de tallo y hojas, una gráfica de series de tiempo, una gráfica de Pareto, un gráfico circular y un gráfico de polígono de frecuencias.
- Determinar cuándo una gráfica es engañosa debido al uso de un eje sin cero o un pictograma que utiliza un objeto de área o volumen para datos unidimensionales.

2-4 Diagramas de dispersión, correlación y regresión

- Desarrollar la capacidad de trazar un diagrama de dispersión de datos pareados.
- Analizar un diagrama de dispersión para determinar si parece haber una correlación entre dos variables.

2-1 Distribuciones de frecuencias para organizar y resumir datos

TABLA 2-2 Tiempos de servicio en el auto para los almuerzos en McDonald's

Tiempo (segundos)	Frecuencia
75-124	11
125-174	24
175-224	10
225-274	3
275-324	2

Concepto clave Cuando se trabaja con grandes conjuntos de datos, una *distribución de frecuencias* (o *tabla de frecuencias*) suele ser útil para la organización y el resumen de los datos. Una distribución de frecuencias nos ayuda a comprender la naturaleza de la *distribución* de un conjunto de datos.

DEFINICIÓN

Una **distribución de frecuencias** (o **tabla de frecuencias**) indica cómo un conjunto de datos se divide en varias categorías (o clases) al listar todas las categorías junto con el número de valores de los datos (frecuencias) que hay en cada una.

Utilicemos los tiempos de servicio de McDonald's para el almuerzo (en segundos) listados en la tabla 2-1. Por otra parte, la tabla 2-2 es una distribución de frecuencias que resume los tiempos de servicio. La **frecuencia** para una clase particular es el número de valores originales que caen en esa clase. Por ejemplo, la primera clase de la tabla 2-2 tiene una frecuencia de 11, por lo que 11 de los tiempos de servicio están entre 75 y 124 segundos, inclusive.

Los siguientes términos estándar se usan con frecuencia para la elaboración de distribuciones y gráficas de frecuencias.

DEFINICIONES

Los **límites inferiores de clase** son las cifras más pequeñas que pueden pertenecer a cada una de las clases. (La tabla 2-2 tiene límites inferiores de clase de 75, 125, 175, 225 y 275).

Los **límites superiores de clase** son las cifras más grandes que pueden pertenecer a cada una de las clases. (La tabla 2-2 tiene límites superiores de clase de 124, 174, 224, 274 y 324).

Las **fronteras de clase** son las cifras que se utilizan para separar las clases, pero sin los espacios creados por los límites de clase. La figura 2-1 muestra los espacios creados por los límites de clase de la tabla 2-2. Se observa que los valores 124.5, 174.5, 224.5 y 274.5 están en los puntos medios de esos espacios. Si se sigue el patrón de las fronteras de clase, puede verse que la frontera de clase inferior es 74.5 y la frontera de clase superior es 324.5. La lista completa de las fronteras de clase es 74.5, 124.5, 174.5, 224.5, 274.5 y 324.5.

Las **marcas de clase** son los valores en el punto medio de las clases. La tabla 2-2 tiene las marcas de clase 99.5, 149.5, 199.5, 249.5 y 299.5. Cada marca de clase se puede encontrar al sumar el límite inferior de clase más el límite superior de clase y dividir la suma por 2.

Anchura de clase es la diferencia entre dos límites inferiores de clase consecutivos (o dos fronteras inferiores de clase consecutivas) en una distribución de frecuencias. La tabla 2-2 usa una anchura de clase de 50. (Los dos primeros límites inferiores de clase son 75 y 125, y su diferencia es 50).

PRECAUCIÓN Encontrar la anchura de clase correcta puede ser complicado. Para la anchura de clase, no cometa el error más común de usar la diferencia entre un límite inferior de clase y un límite superior de clase. Observe en la tabla 2-2 que la anchura de clase es 50, no 49.

PRECAUCIÓN Con respecto a las fronteras de clase, recuerde que dividen la diferencia entre el final de una clase y el comienzo de la siguiente, como se muestra en la figura 2-1.

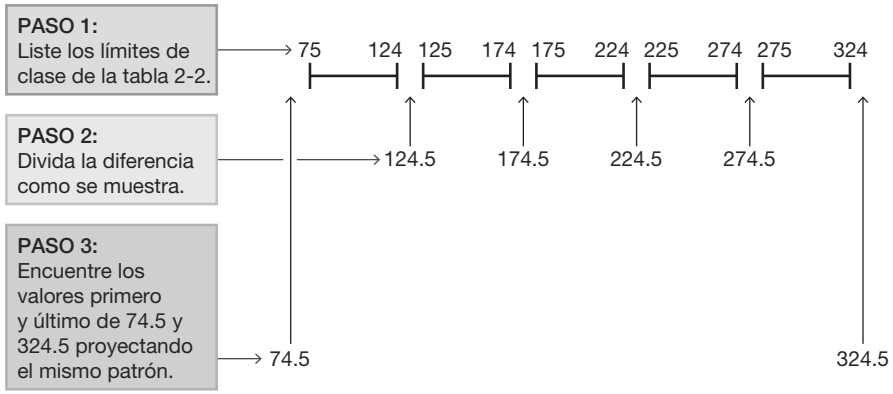


FIGURA 2-1 Determinación de las fronteras de clase a partir de los límites de clase de la tabla 2-2

Procedimiento para elaborar una distribución de frecuencias

Las distribuciones de frecuencias se elaboran para (1) resumir grandes conjuntos de datos, (2) observar la distribución e identificar los valores atípicos, y (3) tener una base para producir gráficas (como los *histogramas* que se introducen en la sección 2-2). Las distribuciones de frecuencias pueden generarse mediante software, pero a continuación se indican los pasos para elaborarlas manualmente:

1. Seleccione el número de clases, normalmente entre 5 y 20. El número de clases puede verse afectado por la conveniencia de utilizar números redondeados. (De acuerdo con la "Regla de Sturges", el número ideal de clases para una distribución de frecuencias puede aproximarse por $1 + (\log n)/(\log 2)$ donde n es el número de valores de los datos. En este libro no seguimos esa regla).
2. Calcule la anchura de clase.

$$\text{Anchura de clase} \approx \frac{(\text{valor máximo de datos}) - (\text{valor mínimo de datos})}{\text{número de clases}}$$

Redondee este resultado para obtener un número conveniente. (Por lo general, es mejor redondear *hacia arriba*). El uso de un número específico de clases no es demasiado importante, y generalmente se recomienda cambiar el número de clases de manera que se tengan valores convenientes para los límites de clase.

3. Elija el valor para el primer límite inferior de clase utilizando el valor mínimo o un valor conveniente por debajo del mínimo.
4. A partir del primer límite inferior de clase y de la anchura de clase, liste los demás límites inferiores de clase. (Esto se hace sumando la anchura de clase al primer límite inferior de clase para obtener el segundo límite inferior de clase. Después se suma la anchura de clase al segundo límite inferior de clase para obtener el tercer límite inferior de clase, y así sucesivamente).

Ni teléfonos ni bañeras

Muchos análisis estadísticos deben considerar las características cambiantes de las poblaciones a través del tiempo. A continuación se presentan algunas observaciones de la vida en Estados Unidos hace 100 años:

- 8% de los hogares tenía un teléfono.
- 14% de los hogares tenía una bañera.
- La esperanza de vida media era de 47 años.
- El salario medio por hora era de 22 centavos.
- Había aproximadamente 230 asesinatos en todo Estados Unidos.

Aunque estas observaciones de hace 100 años indican un marcado contraste con Estados Unidos de hoy, los análisis estadísticos deben considerar siempre características cambiantes de la población que podrían tener efectos más sutiles.

Autores identificados

Entre 1787 y 1788, Alexander Hamilton, John Jay y James Madison publicaron de forma anónima el famoso diario *Federalist Papers*, en un intento por convencer a los neoyorquinos de que deberían ratificar la Constitución. Se conoció la identidad de la mayoría de los autores de los artículos, pero la autoría de 12 de éstos siguió siendo motivo de discusión. Mediante el análisis estadístico de las frecuencias de varias palabras, ahora podemos concluir que *probablemente* James Madison fue el autor de esos 12 documentos. En muchos de los artículos disputados, la evidencia en favor de la autoría de Madison es abrumadora, al grado de que estamos casi seguros de que es lo correcto. Coincidentemente, el autor de este libro ahora vive en una ciudad llamada Madison.

75–
125–
175–
225–
275–

5. Liste los límites inferiores de clase en una columna vertical y después determine e introduzca los límites superiores de clase.
6. Tome cada valor de datos individual y coloque una marca de registro en la clase apropiada. Agregue las marcas de registro para encontrar la frecuencia total de cada clase.

Al elaborar una distribución de frecuencias, asegúrese de que las clases no se superpongan. Cada uno de los valores originales debe pertenecer exactamente a una clase. Incluya todas las clases, incluso aquellas con una frecuencia de cero. Trate de usar la misma anchura para todas las clases, aunque en ocasiones es imposible evitar intervalos abiertos, como "65 años o más".

EJEMPLO 1 Tiempos de servicio para los almuerzos en McDonald's

A partir de los tiempos de servicio para los almuerzos en McDonald's de la tabla 2-1, siga el procedimiento anterior para elaborar la distribución de frecuencias mostrada en la tabla 2-2. Utilice cinco clases.

SOLUCIÓN

Paso 1: Seleccione 5 como el número de clases deseadas.

Paso 2: Calcule la anchura de clase como se muestra a continuación. Observe que redondeamos de 45 a 50, que es un número más cómodo de usar.

$$\text{Anchura de clase} \approx \frac{(\text{valor máximo de datos}) - (\text{valor mínimo de datos})}{\text{número de clases}}$$

$$= \frac{308 - 83}{5} = 45 \approx 50 \text{ (redondeado a un número más conveniente)}$$

Paso 3: El valor mínimo de los datos es 83, que no es un punto de inicio muy conveniente, así que busque un valor más conveniente por debajo de 83, por ejemplo 75, como el primer límite inferior de clase. (Podríamos haber usado 80 o 50 en su lugar).

Paso 4: Sume la anchura de clase de 50 al valor inicial de 75 para obtener el segundo límite inferior de clase de 125. Continúe sumando la anchura de clase de 50 hasta tener cinco límites inferiores de clase. Por lo tanto, los límites inferiores de clase son 75, 125, 175, 225 y 275.

Paso 5: Liste los límites inferiores de clase en forma vertical, como se muestra en el margen. A partir de esta lista, identifique los límites superiores de clase correspondientes como 124, 174, 224, 214 y 324.

Paso 6: Introduzca una marca de registro para cada valor de datos en la clase apropiada. A continuación, sume las marcas de registro para encontrar las frecuencias mostradas en la tabla 2-2.

SU TURNO Resuelva el ejercicio 11 "Old Faithful".

Datos categóricos Hasta ahora hemos analizado distribuciones de frecuencias utilizando sólo conjuntos de datos cuantitativos, pero también es posible utilizar las distribuciones de frecuencias para resumir datos categóricos (o cualitativos, o de atributo), como lo ilustra el ejemplo 2.

EJEMPLO 2 Visitas a las salas de emergencia por lesiones en deportes y actividades recreativas

En la tabla 2-3 se presentan los datos de las siete principales fuentes de lesiones que resultaron en una visita a la sala de urgencias de un hospital en un año reciente (con base en datos de los Centros para el Control de Enfermedades). Los nombres de las actividades son datos categóricos al nivel nominal de medición, pero podemos crear la distribución de frecuencias como se muestra. Resulta sorprendente ver que el ciclismo está en la parte superior de esta lista, pero esto no significa que el ciclismo sea la más peligrosa de las actividades; hay muchas más personas que montan bicicleta que las que juegan fútbol americano, manejan un vehículo todo terreno o hacen cualquiera de las otras actividades listadas.

TABLA 2-3 Visitas anuales a la sala de urgencias por lesiones en deportes y recreación

Actividad	Frecuencia
Ciclismo	26,212
Fútbol americano	25,376
Juegos infantiles	16,706
Baloncesto	13,987
Fútbol	10,436
Béisbol	9,634
Vehículo todo terreno	6,337

En cifras

14: El número de formas de las narices humanas, a partir de un estudio de Abrahim Tamir que fue publicado en el *Journal of Craniofacial Surgery*.

Distribución de frecuencias relativas

Una variación de la distribución de frecuencias básica es una **distribución de frecuencias relativas** o **distribución de frecuencias porcentuales**, en la que cada frecuencia de clase se sustituye por una frecuencia relativa (o proporción) o porcentaje. En este texto usamos el término "distribución de frecuencias relativas", tanto si empleamos frecuencias relativas como porcentajes. Las frecuencias relativas y porcentajes se calculan de la siguiente manera.

$$\text{Frecuencias relativas para una clase} = \frac{\text{frecuencia para una clase}}{\text{suma de todas las frecuencias}}$$

$$\text{Porcentaje para una clase} = \frac{\text{frecuencia para una clase}}{\text{suma de todas las frecuencias}} \times 100\%$$

La tabla 2-4 es un ejemplo de una distribución de frecuencias relativas. Es una variación de la tabla 2-2 en la que cada frecuencia de clase se sustituye por el valor porcentual correspondiente. Debido a que hay 50 valores de datos, se divide cada frecuencia de clase por 50, y después se multiplica por 100%. La primera clase de la tabla 2-2 tiene una frecuencia de 11, así que se divide 11 por 50 para obtener 0.22 y luego se multiplica por 100% para obtener 22%. La suma de los porcentajes debe ser 100%, con una discrepancia mínima permitida para los errores de redondeo, por lo que una suma como 99% o 101% es aceptable. La suma de los porcentajes de la tabla 2-4 es del 100%.

La suma de los porcentajes en una distribución de frecuencias relativas debe ser muy cercana a 100% (con un pequeño margen para errores de redondeo).

TABLA 2-4 Distribución de frecuencias relativas de los tiempos de servicio para el almuerzo en McDonald's

Tiempo (segundos)	Frecuencias relativas
75-124	22%
125-174	48%
175-224	20%
225-274	6%
275-324	4%

Distribución de frecuencias acumuladas

Otra variación de una distribución de frecuencias es una **distribución de frecuencias acumuladas** en la que la frecuencia para cada clase es la suma de las frecuencias para la misma y todas las anteriores. La tabla 2-5 es la distribución de frecuencias acumuladas de la tabla 2-2. A partir de las frecuencias originales de 11, 24, 10, 3, 2, se suma 11 + 24 para obtener la segunda frecuencia acumulada de 35, luego se suma 11 + 24 + 10 para obtener la tercera, y así sucesivamente. Observe en la tabla 2-5 que además del uso de frecuencias acumuladas, los límites de clase son reemplazados por expresiones "menor que", las cuales describen los nuevos intervalos de valores.

TABLA 2-5 Distribución de frecuencias acumuladas de los tiempos de servicio para el almuerzo en McDonald's

Tiempo (segundos)	Frecuencia acumulada
Menor que 125	11
Menor que 175	35
Menor que 225	45
Menor que 275	48
Menor que 325	50

Pensamiento crítico: uso de distribuciones de frecuencias para entender los datos

Al comienzo de esta sección observamos que una distribución de frecuencias puede ayudarnos a entender la *distribución* de un conjunto de datos, que es la naturaleza o forma de la dispersión de los datos sobre el rango de valores (por ejemplo, en forma de campana). Con frecuencia, en estadística es importante determinar si los datos tienen una *distribución nor-*

Gráficas de crecimiento actualizadas

Los pediatras acostumbran utilizar gráficas de crecimiento estandarizadas para comparar el peso y la estatura de sus pacientes con una muestra de otros niños. Se considera que los niños están en un intervalo normal si su peso y estatura caen entre los percentiles 5 y 95. Si están fuera de este intervalo, generalmente se les aplican pruebas para asegurarse de que no tengan problemas médicos de cuidado. Los pediatras ahora son más conscientes de un inconveniente importante de las gráficas: como éstas se basan en niños que vivieron entre 1929 y 1975, las gráficas de crecimiento estaban resultando inexactas. Para rectificar este problema, en 2000 se actualizaron las gráficas para que reflejaran las medidas actuales de millones de niños. Los pesos y las estaturas de los niños son buenos ejemplos de poblaciones que cambian con el paso del tiempo. Esta es la razón de incluir, como un aspecto importante de una población, las características que cambian en los datos con el paso del tiempo.

mal. (Las distribuciones normales se analizan de manera amplia en el capítulo 6.) Los datos que tienen una distribución aproximadamente normal se caracterizan por una distribución de frecuencias con las siguientes características.

Distribución normal

1. Las frecuencias comienzan bajas, luego aumentan a una o dos frecuencias altas y luego disminuyen a una frecuencia baja.
2. La distribución es aproximadamente simétrica: las frecuencias que preceden a la frecuencia máxima deben ser aproximadamente una imagen especular de aquellas que siguen de la frecuencia máxima.

La tabla 2-6 satisface estas dos condiciones. Las frecuencias comienzan en un nivel bajo, aumentan hasta el máximo de 30 y luego disminuyen a una frecuencia baja. Además, las frecuencias de 2 y 8 que preceden al máximo son una imagen especular de las frecuencias 8 y 2 que siguen al máximo. Por lo general, los conjuntos de datos reales no son tan perfectos como en la tabla 2-6, y es necesario usar el juicio para determinar si la distribución está "suficientemente cerca" de satisfacer esas dos condiciones. (Más adelante se presentan otros procedimientos objetivos).

TABLA 2-6 Distribución de frecuencias que muestra una distribución normal

Tiempo	Frecuencia	Distribución normal
75-124	2	← Las frecuencias comienzan bajas, …
125-174	8	
175-224	30	← Aumentan hasta este máximo, …
225-274	8	
275-324	2	← Disminuyen hasta volverse bajas de nuevo.

Análisis de los últimos dígitos En el ejemplo 3 se ilustra este principio:

En ocasiones, las frecuencias de los últimos dígitos revelan cómo se recolectaron o midieron los datos.

TABLA 2-7 Últimos dígitos de los pulsos de la Encuesta Nacional de Salud y Exámenes

Último dígito de los pulsos	Frecuencia
0	455
1	0
2	461
3	0
4	479
5	0
6	425
7	0
8	399
9	0

EJEMPLO 3 Exploración de datos: ¿cómo se midió el pulso?

Al examinar los pulsos medidos a 2219 adultos incluidos en la Encuesta Nacional de Salud y Exámenes, se identifican los últimos dígitos de los pulsos registrados y la distribución de frecuencias para esos últimos dígitos es la presentada en la tabla 2-7. Aquí hay que observar algo: todos los últimos dígitos son números *pares*. Si los pulsos se contaron durante 1 minuto completo, seguramente habría un gran número de ellos que terminen con un dígito *impar*. ¿Entonces qué pasó?

Una explicación razonable es que a pesar de que los pulsos son el número de latidos en 1 minuto, probablemente fueron contados durante 30 segundos y el número de latidos se duplicó. (Las tasas de pulso originales no son todos múltiplos de 4, por lo que podemos descartar un procedimiento de contar durante 15 segundos y luego multiplicar por 4). El análisis de estos últimos dígitos nos revela el método utilizado para obtener los datos.

En muchas encuestas es posible determinar si a los sujetos encuestados se les pidió *reportar* algunos valores, como sus estaturas o pesos, porque desproporcionadamente muchos valores terminan en 0 o 5. Esto es una pista sólida de que el encuestado está redondeando en lugar de ser medido físicamente. ¡Cosas fascinantes!

SU TURNO Resuelva el ejercicio 17 "Análisis de los últimos dígitos".

Brechas En el ejemplo 4 se ilustra este principio:

> **La presencia de brechas puede sugerir que los datos son de dos o más poblaciones diferentes.**

Lo contrario a este principio no es verdadero, porque los datos de diferentes poblaciones no necesariamente producen espacios.

EJEMPLO 4 Exploración de datos: ¿qué nos indica un espacio?

La tabla 2-8 es una distribución de frecuencias de los pesos (en gramos) de monedas seleccionadas al azar. El examen de las frecuencias revela *una gran brecha* entre las monedas de centavo más ligeras y las más pesadas. Esto sugiere que tenemos dos poblaciones distintas: los centavos hechos antes de 1983 son 95% de cobre y 5% de zinc, mientras que los centavos hechos después de 1983 son 2.5% de cobre y 97.5% de zinc, lo que explica la gran diferencia entre los centavos más ligeros y los más pesados representados en la tabla 2-8.

SU TURNO Resuelva el ejercicio 18 "Análisis de los últimos dígitos" y determine si hay una brecha. Si es así, dé una explicación razonable.

TABLA 2-8 Monedas seleccionadas al azar

Peso (en gramos) de los centavos	Frecuencia
2.40-2.49	18
2.50-2.59	19
2.60-2.69	0
2.70-2.79	0
2.80-2.89	0
2.90-2.99	2
3.00-3.09	25
3.10-3.19	8

Comparaciones En el ejemplo 5 se ilustra este principio:

> **La combinación de dos o más distribuciones de frecuencias relativas en una tabla facilita en gran medida las comparaciones entre los datos.**

EJEMPLO 5 Comparación de McDonald's y Dunkin' Donuts

La tabla 2-9 muestra las distribuciones de frecuencias relativas para los tiempos de servicio de los almuerzos servidos en el auto (en segundos) para McDonald's y Dunkin' Donuts. Debido a las grandes diferencias en sus menús, podríamos esperar que los tiempos de servicio fueran muy diferentes. Al comparar las frecuencias relativas en la tabla 2-9, se observa que hay diferencias importantes. Los tiempos de servicio de Dunkin' Donuts parecen ser inferiores a los de McDonald's. Esto no es demasiado sorprendente, dado que probablemente muchas de las órdenes en Dunkin' Donuts consisten en un café y una dona.

TABLA 2-9 Tiempos de servicio para el almuerzo en McDonald's y Dunkin' Donuts

Tiempo (en segundos)	McDonald's	Dunkin' Donuts
25-74		22%
75-124	22%	44%
125-174	48%	28%
175-224	20%	6%
225-274	6%	
275-324	4%	

SU TURNO Resuelva el ejercicio 19 "Ganadores del Oscar".

CENTRO DE TECNOLOGÍA

Distribuciones de frecuencia

Acceda a los complementos tecnológicos, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

A menudo, las distribuciones de frecuencias son fáciles de obtener después de generar un histograma, como se describe en la sección 2-2. Con Statdisk, por ejemplo, es posible generar un histograma con un punto de inicio y una anchura de clase deseados y luego marcar "Bar Labels" para ver la frecuencia de cada clase. Si no se utilizan histogramas, la opción "Sort" (para ordenar los datos) permite ver los valores máximo y mínimo que se utilizan para calcular la anchura de clase. Una vez que se establecen los límites de clase, es fácil encontrar la frecuencia para cada clase usando datos ordenados. Cada paquete de software estadístico incluye una función de ordenación.

2-1 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Tiempos de servicio para las cenas en McDonald's Consulte la tabla anexa que resume los tiempos de servicio (en segundos) de las cenas en McDonald's. ¿Cuántas personas se incluyen en el resumen? ¿Es posible identificar los valores exactos de todos los tiempos de servicio originales?

2. Tiempos de servicio para las cenas en McDonald's Consulte la distribución de frecuencias anexa. ¿Qué problema se crearía al utilizar clases de 60 a 120, 120 a 180, ..., 300 a 360?

3. Distribución de frecuencias relativas Use los porcentajes para elaborar la distribución de frecuencias relativas correspondiente a la distribución de frecuencias anexa de los tiempos de servicio para las cenas en McDonald's.

4. ¿Qué hay de erróneo? Se sabe que las estaturas de los varones adultos tienen una distribución normal, como se describe en esta sección. Un investigador afirma haber seleccionado aleatoriamente a varones adultos y haber medido sus estaturas con la distribución de frecuencia relativa resultante que se muestra aquí. Identifique dos fallas importantes con base en estos resultados.

Tabla para el ejercicio 1

Tiempos de servicio para las cenas en McDonald's

Tiempo (s)	Frecuencia
60-119	7
120-179	22
180-239	14
240-299	2
300-359	5

Tabla para el ejercicio 4

Estatura (cm)	Frecuencia relativa
130-144	23%
145-159	25%
160-174	22%
175-189	27%
190-204	28%

En los ejercicios 5 a 8, identifique la anchura de clase, los puntos medios de clase y los límites de clase para la distribución de frecuencias dada. También identifique el número de individuos incluidos en el resumen. Las distribuciones de frecuencias se basan en datos reales del apéndice B.

5.

Edad (en años) de ganadoras del Oscar como mejor actriz	Frecuencia
20-29	29
30-39	34
40-49	14
50-59	3
60-69	5
70-79	1
80-89	1

6.

Edad (en años) de ganadores del Oscar como mejor actor	Frecuencia
20-29	1
30-39	28
40-49	36
50-59	15
60-69	6
70-79	1

7.

Conteo de plaquetas sanguíneas en hombres	Frecuencia
0-99	1
100-199	51
200-299	90
300-399	10
400-499	0
500-599	0
600-699	1

8.

Conteo de plaquetas sanguíneas en mujeres	Frecuencia
100-199	25
200-299	92
300-399	28
400-499	0
500-599	2

Distribuciones normales. ***En los ejercicios 9 y 10, use una interpretación flexible de los criterios para determinar si una distribución de frecuencias se aproxima a una distribución normal. Dé una explicación breve.***

9. Mejores actrices Considere la distribución de frecuencias del ejercicio 5.

10. Mejores actores Considere la distribución de frecuencias del ejercicio 6.

Elaboración de distribuciones de frecuencias. ***En los ejercicios 11 a 18, utilice los datos indicados para elaborar la distribución de frecuencias. (Los datos de los ejercicios 13 a 16 se pueden descargar de*** **www.pearsonenespañol.com/triola***).*

11. Old Faithful A continuación se listan los tiempos de duración (en segundos) de las erupciones del géiser Old Faithful en el Parque Nacional de Yellowstone. Utilice estos tiempos para elaborar una distribución de frecuencias. Use una anchura de clase de 25 segundos e inicie en un límite inferior de clase de 125 segundos.

125	203	205	221	225	229	233	233	235	236	236	237	238	238	239	240	240
240	240	241	241	242	242	242	243	243	244	245	245	245	245	246	246	248
248	248	249	249	250	251	252	253	253	255	255	256	257	258	262	264	

12. Tornados A continuación se listan las intensidades en la escala F de los tornados recientes en Estados Unidos. Elabore una distribución de frecuencias. ¿Las intensidades parecen tener una distribución normal?

0	4	0	0	1	1	1	0	0	0	1	2	0	1	1	0	1	0	1	1	1	1	0
0	1	0	0	1	0	0	1	1	1	3	0	0	0	2	0	3	0	0	0	0	0	

13. Tiempos de servicio para almuerzos en Burger King Consulte el conjunto de datos 25 "Comida rápida" y utilice los tiempos de servicio en auto para los almuerzos en Burger King. Inicie en un límite inferior de clase de 70 segundos y utilice una anchura de clase de 40 segundos.

14. Tiempos de servicio para cenas en Burger King Consulte el conjunto de datos 25 "Comida rápida" y utilice los tiempos de servicio en auto para las cenas en Burger King. Inicie en un límite inferior de clase de 30 segundos y utilice una anchura de clase de 40 segundos.

15. Tiempos de servicio para almuerzos en Wendy's Consulte el conjunto de datos 25 "Comida rápida" y utilice los tiempos de servicio en auto para los almuerzos en Wendy's. Inicie en un límite inferior de clase de 70 segundos y utilice una anchura de clase de 80 segundos. ¿Parece que la distribución es una distribución normal?

16. Tiempos de servicio para cenas en Wendy's Consulte el conjunto de datos 25 "Comida rápida" y use los tiempos de servicio para las cenas en Wendy's. Inicie en un límite inferior de clase de 30 segundos y utilice una anchura de clase de 40 segundos. Use una interpretación flexible de una distribución normal y diga si esta distribución parece ser una distribución normal.

17. Análisis de los últimos dígitos El autor obtuvo las estaturas de sus estudiantes de estadística como parte de un experimento realizado en clase. A continuación se listan los últimos dígitos de esas estaturas. Elabore una distribución de frecuencias con 10 clases. Con base en la distribución, ¿parece que las estaturas fueron reportadas o realmente medidas? ¿Qué puede saber usted de la exactitud de los resultados?

0	0	0	0	0	0	0	0	0	1	1	2	3	3	3	4	5	5	5	
5	5	5	5	5	5	5	5	5	5	5	5	6	6	8	8	8	9		

18. Análisis de los últimos dígitos Los pesos de los encuestados se registraron como parte de la Entrevista del Sondeo de la Salud en California. A continuación se listan los últimos dígitos de los pesos de 50 encuestados seleccionados al azar. Elabore una distribución de frecuencias con 10 clases. Con base en la distribución, ¿parece que los pesos son reportados o realmente medidos? ¿Qué puede saber usted de la exactitud de los resultados?

5	0	1	0	2	0	5	0	5	0	3	8	5	0	5	0	5	6	0	0	0	0	0	0	8
5	5	0	4	5	0	0	4	0	0	0	0	0	8	0	9	5	3	0	5	0	0	0	5	8

Frecuencias relativas para comparaciones. ***En los ejercicios 19 y 20, elabore las distribuciones de frecuencias relativas y responda las preguntas.***

19. Ganadores del Oscar Elabore una tabla (similar a la tabla 2-9 de la página 47) que incluya frecuencias relativas basadas en las distribuciones de frecuencia de los ejercicios 5 y 6, y después compare las edades de las actrices y los actores ganadoras del Oscar. ¿Hay diferencias notables?

20. Conteos de plaquetas sanguíneas Elabore una tabla (similar a la tabla 2-9 de la página 47) que incluya frecuencias relativas basadas en las distribuciones de frecuencia de los ejercicios 7 y 8, y después compárelas. ¿Hay diferencias notables?

Distribuciones de frecuencia acumuladas. ***En los ejercicios 21 y 22, elabore la distribución de frecuencias acumuladas correspondiente a la distribución de frecuencias del ejercicio indicado.***

21. Ejercicio 5 (Edad de la ganadora del Oscar a la mejor actriz).

22. Ejercicio 6 (Edad del ganador al Oscar al mejor actor).

Datos categóricos. ***En los ejercicios 23 y 24, utilice los datos categóricos dados para elaborar la distribución de frecuencias relativas.***

23. Ensayo clínico Cuando se administró XELJANZ (tofacitinib) como parte de un ensayo clínico para un tratamiento de artritis reumatoide, 1336 sujetos recibieron dosis de 5 mg del fármaco, y a continuación se muestra el número de reacciones adversas: 57 tuvieron dolores de cabeza, 2l tuvieron hipertensión, 60 infecciones del tracto respiratorio, 51 nasofaringitis y 53 diarrea. ¿Alguna de estas reacciones adversas parece ser mucho más común que las otras? (*Sugerencia:* Encuentre las frecuencias relativas usando solamente las reacciones adversas, no el número total de sujetos tratados).

24. Nacimientos Los nacimientos naturales seleccionados aleatoriamente de cuatro hospitales en el Estado de Nueva York ocurrieron en los días de la semana (en orden de lunes a domingo) con las siguientes frecuencias: 52, 66, 72, 57, 57, 43, 53. ¿Parece que los nacimientos ocurren en los días de la semana con la misma frecuencia?

Conjuntos grandes de datos. ***Los ejercicios 25 a 28 implican grandes conjuntos de datos, por lo que debe usarse la tecnología. Las listas completas de los datos no se incluyen en el apéndice B, pero pueden descargarse del sitio web*** **www.pearsonenespañol.com/triola.** ***Utilice los datos indicados y elabore la distribución de frecuencias.***

25. Presión arterial sistólica Utilice la presión arterial sistólica de 300 sujetos incluida en el conjunto de datos 1 "Datos corporales". Use una anchura de clase de 20 mm Hg e inicie en un límite inferior de clase de 80 mm Hg. ¿Parece que la distribución de frecuencias es una distribución normal?

26. Presión arterial diastólica Utilice la presión arterial diastólica de 300 sujetos incluida en el conjunto de datos 1 "Datos corporales". Use una anchura de clase de 15 mm Hg e inicie con un límite inferior de clase de 40 mm Hg. ¿Parece que la distribución de frecuencias es una distribución normal?

27. Magnitudes de terremoto Utilice las magnitudes de 600 terremotos incluidas en el conjunto de datos 21 "Terremotos". Use una anchura de clase de 0.5 e inicie con un límite inferior de clase de 1.00. ¿Parece que la distribución de frecuencias es una distribución normal?

28. Profundidades de terremoto Utilice las profundidades (en km) de 600 terremotos incluidas en el conjunto de datos 21 "Terremotos". Use una anchura de clase de 10.0 km e inicie con un límite inferior de clase de 0.0 km. ¿Parece que la distribución de frecuencias es una distribución normal?

2-1 Más allá de lo básico

29. Interpretación de efectos de los valores atípicos Consulte, en el conjunto de datos 30 "Latas de aluminio" del apéndice B, las cargas axiales de latas de aluminio de 0.0111 pulgadas de grosor. Una carga axial es la fuerza a la cual la parte superior de una lata se colapsa. La carga de 504 lb es un valor atípico porque está muy alejado de todos los demás valores. Elabore una distribución de frecuencias que incluya el valor de 504 lb, y después otra distribución de frecuencias sin el valor de 504 libras. En ambos casos, inicie la primera clase en 200 libras y utilice una anchura de clase de 20 lb. Indique una generalización sobre el efecto de un valor atípico en una distribución de frecuencias.

2-2 Histogramas

PARTE 1 Conceptos básicos de los histogramas

Concepto clave Una distribución de frecuencias es una herramienta útil para resumir datos e investigar su distribución; una herramienta incluso mejor es un *histograma*, una gráfica más fácil de interpretar que una tabla de números.

DEFINICIÓN

Un **histograma** es una gráfica que consiste en barras adyacentes de igual anchura dibujadas (a menos que haya espacios en los datos). La escala horizontal representa clases de valores cuantitativos, y la escala vertical representa sus frecuencias. Las alturas de las barras corresponden a los valores de frecuencia.

Usos importantes de un histograma

- Despliega visualmente la forma de la *distribución* de los datos.
- Muestra la ubicación del *centro* de los datos.
- Muestra la *dispersión* de los datos.
- Identifica los valores *atípicos*.

Un histograma es en esencia una gráfica de una distribución de frecuencias. Por ejemplo, la figura 2-2 muestra el histograma generado por Minitab correspondiente a la distribución de frecuencias dada en la tabla 2-2 de la página 42.

Las frecuencias de clase deben usarse para la escala vertical, la cual se debe etiquetar como en la figura 2-2. No hay un acuerdo universal sobre el procedimiento para seleccionar los valores que se usarán para ubicar las barras a lo largo de la escala horizontal, pero es común utilizar los límites de clase (como se muestra en la figura 2-2) o los puntos medios de clase o los límites de clase o algún otro valor. Con frecuencia, resulta más sencillo usar los puntos medios de clase para la escala horizontal. Por lo general, los histogramas se pueden generar usando software.

Histograma de frecuencias relativas

Un **histograma de frecuencias relativas** tiene la misma forma y escala horizontal que un histograma, pero la escala vertical utiliza frecuencias relativas (como porcentajes o proporciones) en vez de frecuencias reales. La figura 2-3 es el histograma de frecuencias relativas correspondiente a la figura 2-2.

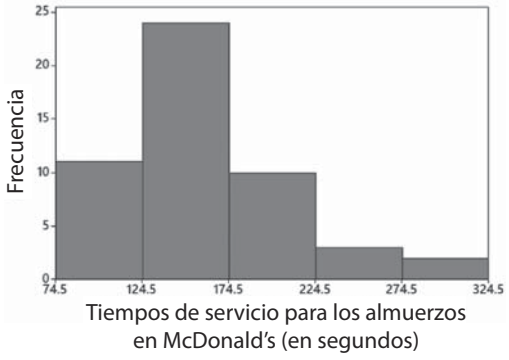


FIGURA 2-2 Histograma de frecuencias relativas del tiempo de servicio para los almuerzos en auto en McDonald's (en segundos)

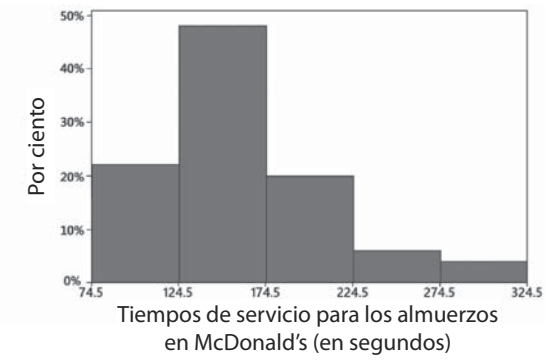


FIGURA 2-3 Histograma del tiempo de servicio para los almuerzos en auto en McDonald's (en segundos)

Pensamiento crítico: Interpretación de histogramas

Aunque la creación de histogramas es lo más divertido que los seres humanos pueden hacer, el objetivo final es *entender* las características de los datos. Se pueden explorar datos analizando el histograma para ver qué se puede aprender de ellos con "CVDOT": el centro de los datos, la variación (que se analizará de manera extensa en la sección 3-2), la forma de la distribución, si hay valores atípicos (*outliers*, en inglés)(muy alejados de los demás valores) y el tiempo (si hay algún cambio en las características de los datos a lo largo del tiempo). Al examinar la figura 2-2, se observa que el histograma está centrado alrededor de 160 o 170 segundos, los valores varían desde alrededor de 75 segundos hasta 325, y la distribución es *muy aproximada* a una forma de campana. No hay valores atípicos y cualquier cambio en el tiempo es irrelevante para estos datos.

Formas comunes de distribución

Los histogramas que se muestran en la figura 2-4 representan cuatro formas comunes de distribución.

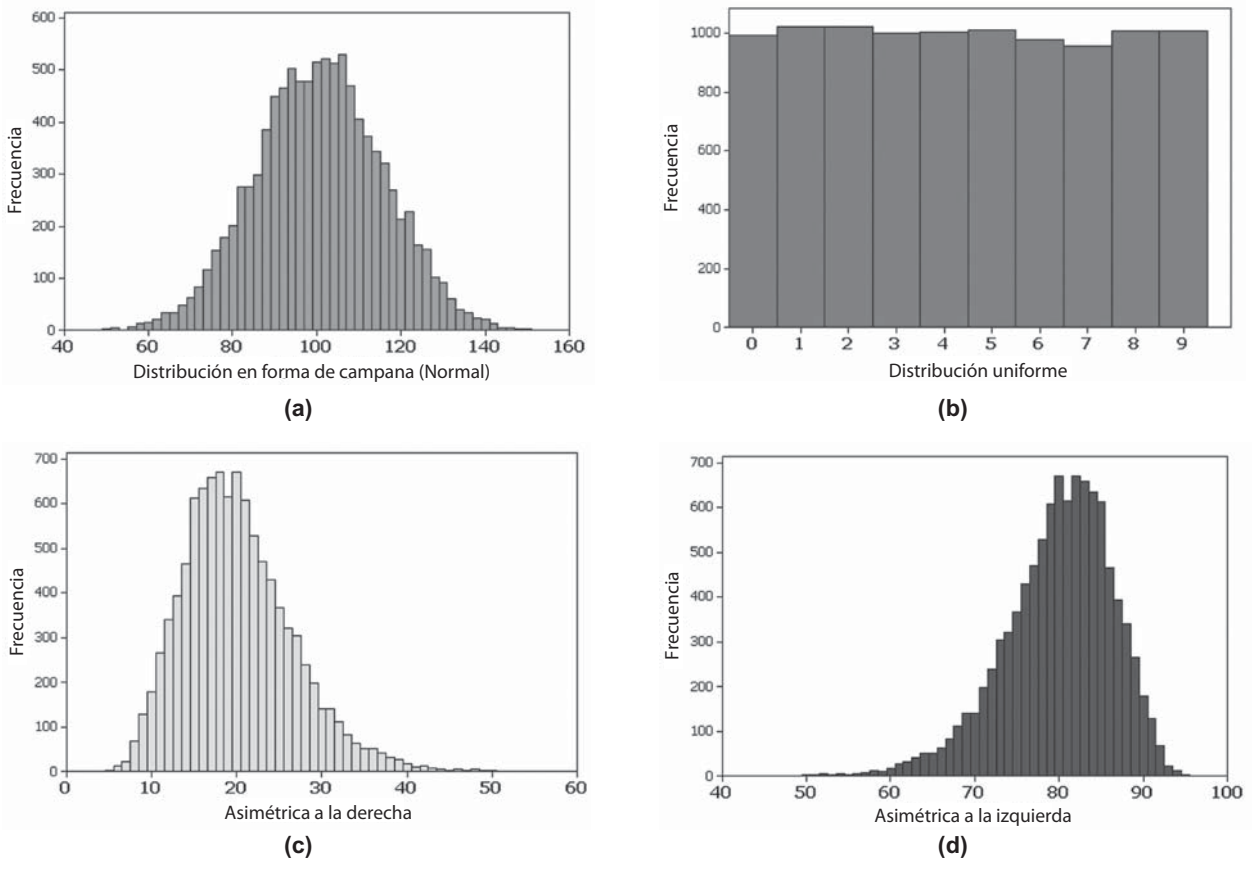


FIGURA 2-4 Distribuciones comunes

Distribución normal

Cuando se grafica como un histograma, una distribución normal tiene una forma de "campana" similar a la superpuesta en la figura 2-5. Muchos métodos estadísticos requieren que los datos muestrales provengan de una población con distribución aproximadamente normal, y a menudo se puede usar un histograma para juzgar si este requisito se cumple (aunque existen métodos más avanzados y menos subjetivos para determinar si la distribución es una distribución normal). Las gráficas cuantilares normales son muy útiles para evaluar la normalidad: vea la parte 2 de esta sección.

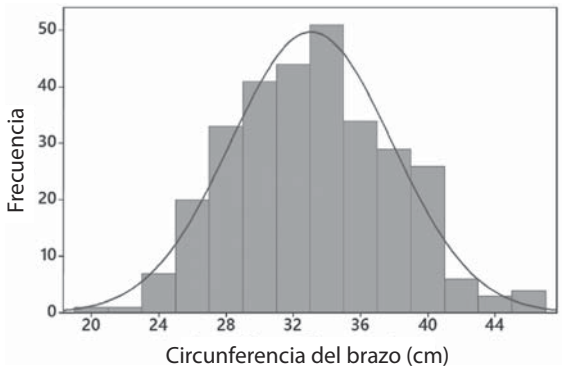


FIGURA 2-5 Distribución en forma de campana de las circunferencias del brazo
Debido a que este histograma tiene una forma aproximada de campana, se dice que los datos tienen una *distribución normal*. (En el capítulo 6 se dará una definición más rigurosa).

Distribución uniforme

Los diferentes valores posibles se producen con aproximadamente la misma frecuencia, por lo que las alturas de las barras en el histograma son aproximadamente uniformes, como en la figura 2-4(b) que muestra los resultados de los dígitos de las loterías estatales.

Asimetría

Una distribución de datos es asimétrica si se extiende más hacia un lado que hacia el otro. Los datos **asimétricos a la derecha** (también llamados *positivamente asimétricos*) tienen una cola derecha más larga, como en la figura 2-4(c). Los ingresos anuales de los estadounidenses adultos son positivamente asimétricos. Los datos **asimétricos a la izquierda** (también llamados *negativamente asimétricos*) tienen una cola izquierda más larga, como en la figura 2-4(d). Los datos de la duración de la vida en los seres humanos son asimétricos a la izquierda. (Aquí hay un truco de mnemotecnia para recordar la asimetría: una distribución asimétrica a la derecha se parece a los dedos del pie derecho, y una asimétrica a la izquierda se parece a los dedos del pie izquierdo). Las distribuciones asimétricas a la derecha son más comunes que las asimétricas a la izquierda porque a menudo es más fácil obtener valores excepcionalmente grandes que valores muy pequeños. Con los ingresos anuales, por ejemplo, es imposible obtener valores por debajo de cero, pero hay algunas personas que ganan millones o miles de millones de dólares al año. Por lo tanto, los ingresos anuales tienden a ser asimétricos a la derecha.

En cifras

2.5 trillones de bytes: Cantidad de datos que generamos el año pasado. (Un trillón es un 1 seguido por 18 ceros).

Para recordar la asimetría:

Asimetría a la izquierda:	**Se asemeja a los dedos del pie izquierdo**
Asimetría a la derecha:	**Se asemeja a los dedos del pie derecho**

PARTE 2 Evaluación de la normalidad con gráficas cuantilares normales

Algunos métodos realmente importantes que se presentan en capítulos subsecuentes tienen el requisito de que los datos muestrales provengan de una población con distribución normal. Los histogramas pueden ser útiles para determinar si se satisface tal requisito de normalidad, pero no son muy útiles con conjuntos de datos pequeños. En la sección 6-5 se analizan métodos para *evaluar la normalidad*, es decir, para determinar si los datos muestrales provienen de una población normalmente distribuida. Dicha sección incluye un procedimiento para trazar *gráficas cuantilares normales*, que son fáciles de generar usando tecnologías como Statdisk, Minitab, XLSTAT, StatCrunch o una calculadora TI-83/84 Plus. La interpretación de una gráfica cuantilar normales se basa en los siguientes criterios:

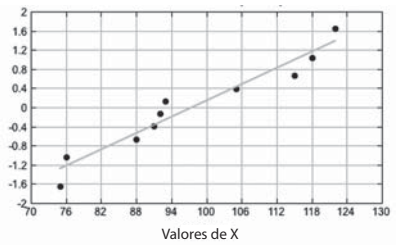
Criterios para evaluar la normalidad con una gráfica cuantilar normal

Distribución normal: La distribución de una población es *normal* si el patrón de los puntos en la gráfica cuantilar normal está razonablemente cerca de una línea recta y los puntos no muestran un patrón sistemático diferente a un patrón lineal.

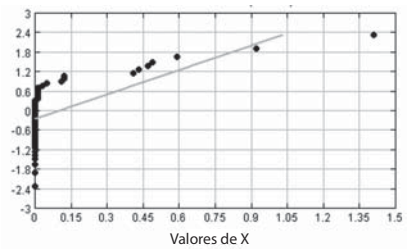
Distribución no normal: La distribución de la población *no* es normal si la gráfica cuantilar normal cumple una o dos de las siguientes condiciones:

- Los puntos no se encuentran razonablemente cerca de un patrón de línea recta.
- Los puntos muestran algún *patrón sistemático* diferente a un patrón de línea recta.

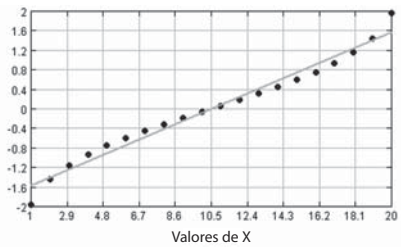
Los siguientes son ejemplos de gráficas cuantilares normales. Los procedimientos para crear tales gráficas se describen en la sección 6-5.



Distribución normal: Los puntos están razonablemente cerca de un patrón lineal, y no hay otro patrón sistemático diferente.



Distribución no normal: Los puntos no se encuentran razonablemente cerca de una línea recta.



Distribución no normal: Los puntos muestran un patrón sistemático que no es un patrón de línea recta.

CENTRO DE TECNOLOGÍA

Histogramas

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Statdisk

1. Haga clic en **Data** del menú superior.
2. Seleccione **Histogram** en el menú desplegable.
3. Seleccione la columna de datos deseada.
4. Haga clic en **Plot**.
5. Marque **Bar Labels** en *Plot Options* para ver la frecuencia de cada clase.
6. Marque **User Defined** en *Plot Options* para utilizar su propia anchura de clase y punto de inicio.

Sugerencia: Este procedimiento también es una manera fácil de identificar frecuencias en una distribución de frecuencias.

Minitab

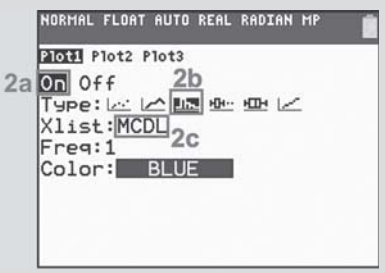
1. Haga clic en **Graph** del menú superior.
2. Seleccione **Histogram** en el menú desplegable.
3. Seleccione el histograma **Simple** y haga clic en **OK**.
4. Haga clic en la columna de datos deseada, luego haga clic en **Select** y después en **OK**.
5. Cambie la anchura de clase predeterminada y el punto de inicio según sea necesario haciendo clic con el botón derecho en el eje horizontal y seleccionando **Edit X Scale**.
   - Seleccione la pestaña **Scale** para introducir la ubicación de las marcas de escala.
   - Seleccione la pestaña **Binning** para introducir los puntos medios de clase.

StatCrunch

1. Haga clic en **Graph** del menú superior.
2. Seleccione **Histogram** en el menú desplegable.
3. Seleccione la columna de datos deseada.
4. Para personalizar el histograma, introduzca el punto de inicio deseado y la anchura de clase en **Bins**.
5. Haga clic en **Compute!**

Calculadora TI-83/84 Plus

1. Abra el menú **STAT PLOTS** pulsando 2ND, Y=.
2. Presione ENTER para acceder a la pantalla de configuración *Plot 1* como se muestra:
   a. Seleccione **ON** y presione ENTER.
   b. Seleccione la opción de la gráfica de barras y presione **ENTER**.
   c. Introduzca el nombre de la lista que contiene los datos.
3. Presione ZOOM, luego 9 (ZoomStat) para generar el histograma predeterminado.
4. Presione TRACE y use ◄ ► para ver las fronteras de clase y las frecuencias de cada clase.
5. Presione WINDOW para personalizar la anchura y las fronteras de clase. Presione GRAPH para ver el histograma.



CENTRO DE TECNOLOGÍA *continuación*

Histogramas

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

Excel

Es extremadamente difícil generar histogramas en Excel; se debe utilizar el complemento XLSTAT:

1. Seleccione la pestaña **XLSTAT** en la cinta de opciones.
2. Haga clic en el botón **Visualizing Data**.
3. Seleccione **Histograms** en el menú desplegable.
4. Introduzca el rango de celdas que contiene los datos deseados. Haga clic en **Sample labels** si la primera celda contiene un nombre de datos.
5. Haga clic en **OK** para generar un histograma predeterminado.

Sugerencia: Para personalizar, introduzca las fronteras de clase deseadas en una columna, seleccione la pestaña **Options**, haga clic en **User Defined** e ingrese el rango de celdas que contiene las fronteras en el cuadro.

2-2 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Estaturas Las estaturas de los varones adultos se distribuyen normalmente. Si se selecciona aleatoriamente una gran muestra de estaturas de varones adultos y las estaturas se ilustran en un histograma, ¿cuál será la forma de ese histograma?

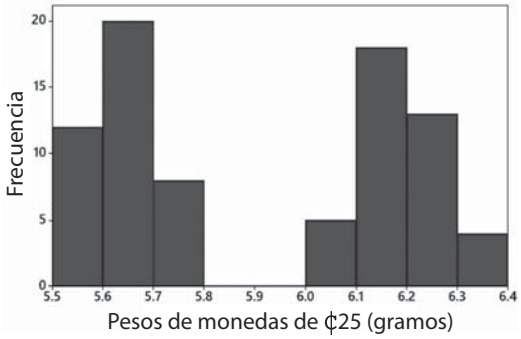
2. Más estaturas La población de las estaturas de los varones adultos se distribuye normalmente. Si obtenemos una muestra de respuesta voluntaria de 5000 de esas estaturas, ¿un histograma de los datos muestrales tendrá forma de campana?

3. Conteo de plaquetas sanguíneas A continuación se listan conteos de plaquetas sanguíneas (1000 células/μl) seleccionados aleatoriamente en adultos de Estados Unidos. ¿Por qué no tiene sentido elaborar un histograma para este conjunto de datos?

191 286 263 193 193 215 162 646 250 386

4. Conteo de plaquetas sanguíneas Si se recolecta una muestra de conteos de plaquetas sanguíneas mucho mayor que la muestra del ejercicio 3, y si dicha muestra incluye un único valor atípico, ¿cómo aparecerá ese valor atípico en un histograma?

Interpretación de un histograma. ***En los ejercicios 5 a 8, responda las preguntas con referencia al siguiente histograma generado en Minitab, que representa los pesos (en gramos) de monedas de veinticinco centavos listados en el conjunto de datos 29 "Pesos de monedas" del apéndice B (los gramos son en realidad unidades de masa y los valores mostrados en la escala horizontal están redondeados).***



5. Tamaño de muestra ¿Cuál es el número aproximado de monedas de veinticinco centavos representado en las tres barras que se encuentran más a la izquierda?

6. Anchura de clase y límites de clase Indique los valores aproximados de la anchura de clase y los límites de clase inferior y superior de la clase representada en la barra más alejada hacia la izquierda.

7. Histograma de frecuencias relativas ¿Cómo cambiaría la forma del histograma si la escala vertical utilizara frecuencias relativas expresadas en porcentajes en vez de los conteos de frecuencias reales que se muestran aquí?

8. Espacio ¿Cuál es una explicación razonable para el espacio entre las monedas con pesos de 5.5 gramos a 5.8 gramos y el grupo de monedas con pesos de 6.0 gramos y 6.4 gramos? (*Sugerencia:* Considere las columnas de las monedas de 25 centavos en el conjunto de datos 29 "Pesos de monedas" del apéndice B).

Elaboración de histogramas. ***En los ejercicios 9 a 16, trace los histogramas y responda las preguntas dadas.***

9. Old Faithful Utilice la distribución de frecuencias del ejercicio 11 en la sección 2-1 de la página 49 para elaborar un histograma. ¿La gráfica parece ser de una población con distribución normal?

10. Tornados Utilice la distribución de frecuencias del ejercicio 12 en la sección 2-1 de la página 49 para elaborar un histograma. ¿El histograma parece ser asimétrico? Si es así, identifique el tipo de asimetría.

11. Tiempos de servicio para el almuerzo en Burger King Utilice la distribución de frecuencias del ejercicio 13 en la sección 2-1 de la página 49 para elaborar un histograma. ¿El histograma parece ser asimétrico? Si es así, identifique el tipo de asimetría.

12. Tiempos de servicio para la cena en Burger King Utilice la distribución de frecuencias del ejercicio 1.1 en la sección 2-1 de la página 49 para elaborar un histograma. Utilice una interpretación estricta de los criterios para una distribución normal, ¿parece que el histograma representa los datos de una población con distribución normal?

13. Tiempos de servicio para el almuerzo en Wendy's Utilice la distribución de frecuencias del ejercicio 15 en la sección 2-1 de la página 49 para elaborar un histograma. ¿El histograma parece ser asimétrico? Si es así, identifique el tipo de asimetría.

14. Tiempos de servicio para la cena en Wendy's Utilice la distribución de frecuencias del ejercicio 16 en la sección 2-1 de la página 49 para elaborar un histograma. Utilice una interpretación estricta de los criterios para una distribución normal, ¿parece que el histograma representa datos de una población con distribución normal?

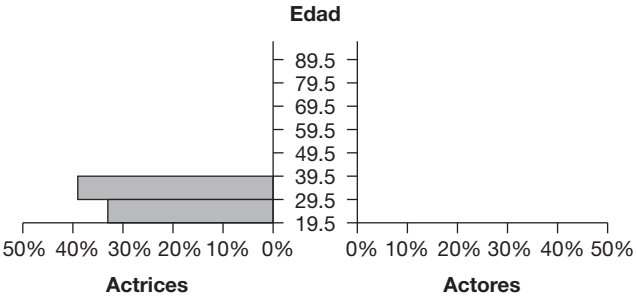
15. Análisis de los últimos dígitos Utilice la distribución de frecuencias del ejercicio 17 en la sección 2-1 de la página 49 para elaborar un histograma. ¿Qué puede deducirse de la distribución de los dígitos? Específicamente, ¿las estaturas parecen ser reportadas o realmente medidas?

16. Análisis de los últimos dígitos Utilice la distribución de frecuencias del ejercicio 18 en la sección 2-1 de la página 49 para elaborar un histograma. ¿Qué puede deducirse de la distribución de los dígitos? Específicamente, ¿las estaturas parecen ser reportadas o realmente medidas?

2-2 Más allá de lo básico

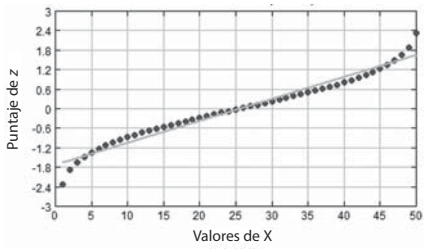
17. Histogramas de frecuencias relativas espalda con espalda Cuando se usan histogramas para comparar dos conjuntos de datos, en ocasiones es difícil hacer comparaciones con ellos. Un histograma de frecuencias relativas espalda con espalda tiene un formato que facilita en gran medida la comparación. En lugar de frecuencias, se deben usar frecuencias relativas (porcentajes o proporciones) para que las comparaciones no sean difíciles cuando hay diferentes tamaños de muestra. Utilice las distribuciones de frecuencias relativas de las edades de las actrices y los actores ganadores del Oscar en el ejercicio 19 de la sección 2-1 en la página 49 y complete los histogramas de frecuencias relativas

espalda con espalda que se muestran a continuación. Después, utilice el resultado para comparar los dos conjuntos de datos.

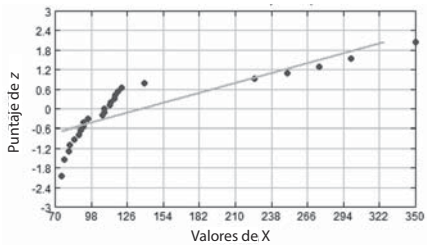


18. Interpretación de las gráficas cuantilares normales ¿Cuáles de las siguientes gráficas cuantilares normales parecen representar datos de una población con distribución normal? Explique.

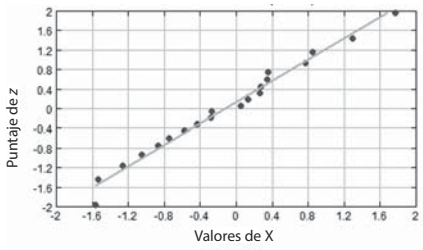
a.



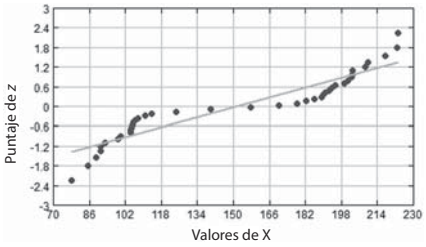
b.



c.



d.



2-3 Gráficas que informan y gráficas que engañan

Concepto clave En la sección 2-2 se introdujo el histograma, y esta sección presenta otras gráficas comunes que fomentan la comprensión de los datos. También se estudian algunas gráficas inexactas porque crean impresiones sobre los datos que, de alguna manera, son engañosas o erróneas.

La era de las gráficas encantadoras y primitivas dibujadas a mano ha pasado a la historia, y ahora la tecnología proporciona herramientas poderosas para generar una gran variedad de gráficas. ¡Aquí vamos!

Gráficas que informan

Gráficas de puntos

Una **gráfica de puntos** consiste en un gráfico de datos *cuantitativos* en el que cada valor de datos se representa como un punto sobre una escala horizontal de valores. Los puntos que representan valores iguales se apilan.

Características de una gráfica de puntos

- Muestra la forma de la distribución de los datos.
- Por lo general, es posible recrear la lista original de datos.

El poder de una gráfica

Con ventas anuales cercanas a los $13 mil millones y con alrededor de 50 millones de usuarios, el fármaco Lipitor (nombre genérico, atorvastatina) de Pfizer se ha convertido en el medicamento de prescripción más redituable y más utilizado de la historia. Al inicio de su desarrollo, Lipitor se comparó con otros fármacos (Zocor [simvastatina], Mevacor [lovastatina], Lescol [fluvastatina] y Pravachol [pravastatina]) en un proceso que implicó ensayos controlados. El resumen del informe incluyó una gráfica que mostraba una curva del Lipitor con un incremento más pronunciado que las curvas de los otros medicamentos, lo cual demostraba visualmente que Lipitor era más eficaz para reducir el colesterol que los otros fármacos. Pat Kelly, que en ese entonces era ejecutiva de marketing de nivel superior en Pfizer, declaró: "Nunca olvidaré cuando vi esa gráfica [...] En ese momento dije '¡Ah!, ahora sé de qué se trata'. ¡Podemos comunicar esto!". La Administración de Alimentos y Medicamentos (FDA) de Estados Unidos aprobó el Lipitor y permitió a Pfizer incluir la gráfica con cada prescripción. El personal de ventas de la empresa también distribuyó la gráfica entre los médicos.

EJEMPLO 1 Gráfica de puntos de los pulsos en varones

La figura 2-6 muestra una gráfica de puntos de los pulsos (pulsaciones por minuto) de los varones del conjunto de datos 1 "Datos corporales" del apéndice B. Los dos puntos apilados encima de la posición 50 indican que dos de los pulsos son de 50. (En esta gráfica de puntos, la escala horizontal permite sólo números pares, pero los pulsos originales son todos pares).

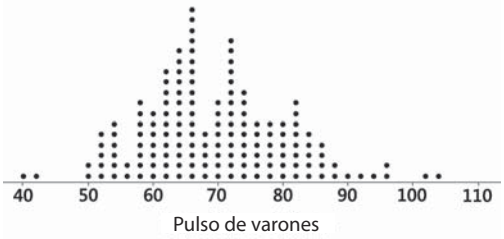


FIGURA 2-6 Gráfica de puntos de los pulsos de varones

SU TURNO Resuelva el ejercicio 5 "Pulsos".

Diagramas de tallo y hojas

Un **diagrama de tallo y hojas** representa datos *cuantitativos* separando cada valor en dos partes: el tallo (por ejemplo el dígito más a la izquierda) y las hojas (como el dígito más a la derecha). A menudo, los mejores diagramas tallo y hojas se obtienen al redondear primero los valores de los datos originales. Además, los diagramas de tallo y hojas se pueden *expandir* para incluir más filas, o bien *condensar* para incluir menos, como en el ejercicio 21 "Diagramas de tallo y hojas expandidos".

Características de un diagrama de tallo y hojas

- Muestra la forma de la distribución de los datos.
- Conserva los valores de los datos originales.
- Los datos muestrales aparecen ordenados.

EJEMPLO 2 Diagrama de tallo y hojas de los pulsos de varones

El siguiente diagrama muestra los pulsos de los varones en el conjunto de datos 1 "Datos corporales" del apéndice B. El pulso más bajo de 40 se separa en el tallo 4 y la hoja 0. Los tallos y las hojas se ordenan de manera creciente, no el orden establecido por la lista original. Si el tallo se coloca en forma horizontal, es posible ver la distribución de los pulsos de la misma manera que se vería en un histograma o una gráfica de puntos.

```
4 | 02  <------------------ Los pulsos son 40 y 42
5 | 00222224444446688888888
6 | 0000000222222222224444444444444666666666666666666688888
7 | 00000000222222222222224444444444666666888888
8 | 0000002222222244444666688
9 | 02466  <------------------ Los pulsos son 90, 92, 94, 96, 96
10| 24
```

SU TURNO Resuelva el ejercicio 7 "Pulsos".

Gráfica de series de tiempo

Una **gráfica de series de tiempo** es una gráfica de *datos dados en series* de tiempo, los cuales son datos cuantitativos recopilados en diferentes momentos, por ejemplo cada mes o cada año.

Característica de una gráfica de series de tiempo

- Revela información sobre tendencias a través del tiempo.

EJEMPLO 3 **Gráfica de series de tiempo de bajas de agentes de la ley**

La gráfica de series de tiempo que se muestra en la figura 2-7 presenta el número anual de muertes de agentes de la ley en Estados Unidos. Vea que ocurrió un pico en 2001, el año de los ataques terroristas del 11 de septiembre. Con excepción de los datos de 2001, parece haber una ligera tendencia a la baja.

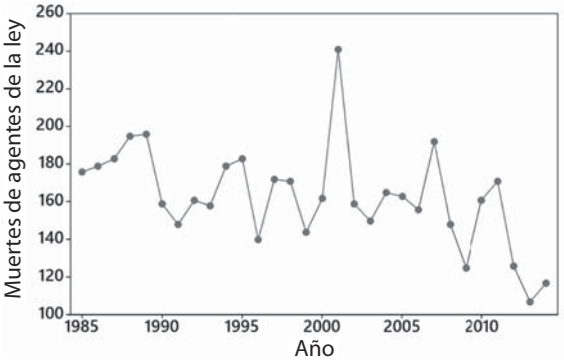


FIGURA 2-7 **Gráfica de series de tiempo de las bajas de agentes de la ley**

SU TURNO Resuelva el ejercicio 9 "Brecha de género en el salario".

Florence Nightingale

Florence Nightingale (1820-1910) es reconocida como la fundadora de la profesión de la enfermería, aunque salvó miles de vidas también utilizando la estadística. Cuando encontraba un hospital insalubre y con desabasto, mejoraba tales condiciones y después utilizaba la estadística para convencer a otros de la necesidad de una reforma médica de mayor alcance. Diseñó gráficas originales para ilustrar que, durante la Guerra de Crimea, murieron más soldados como resultado de las condiciones insalubres que en combate. Florence Nightingale fue pionera en el uso de la estadística social y de las técnicas de gráficas.

Gráficas de barras

Una **gráfica de barras** utiliza barras de igual anchura para mostrar las frecuencias de las categorías de datos *categóricos* (o cualitativos). Las barras pueden o no estar separadas por pequeños espacios.

Característica de una gráfica de barras

- Muestra la distribución relativa de datos categóricos para que sea más fácil comparar las diferentes categorías.

Gráficos de Pareto

Una **gráfica de Pareto** es una gráfica de barras para datos categóricos, con la estipulación añadida de que *las barras se ordenan de manera descendente* de acuerdo con las frecuencias; por ello las barras disminuyen de altura de izquierda a derecha.

Características de una gráfica de Pareto

- Muestra la distribución relativa de datos categóricos para que sea más fácil comparar las diferentes categorías.
- Orienta la atención hacia las categorías más importantes.

EJEMPLO 4 **Gráfica de Pareto para los robos de vehículos acuáticos**

La figura 2-8 muestra los tipos de vehículos acuáticos más robados en un año reciente. Se puede ver que los robos de motos acuáticas son el problema más grave.

continúa

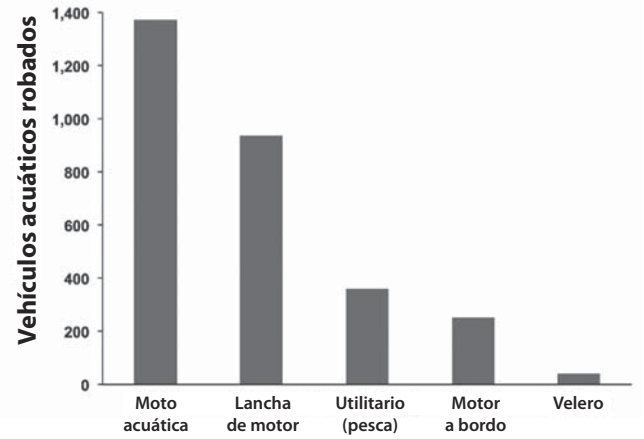


FIGURA 2-8 Gráfica de Pareto de vehículos acuáticos robados

SU TURNO Resuelva el ejercicio 11 "Rectificaciones en revistas".

Gráficas circulares

Una **gráfica circular** es una gráfica muy común que representa datos categóricos como rebanadas de un círculo; el tamaño de cada rebanada es proporcional al conteo de frecuencias para la categoría. Aunque las gráficas circulares son muy comunes, no son tan efectivas como las gráficas de Pareto.

Característica de una gráfica circular

- Muestra la distribución de datos categóricos en un formato de uso común.

EJEMPLO 5 Gráfica circular de vehículos acuáticos robados

La figura 2-9 es una gráfica circular con los mismos datos de robo del ejemplo 4. El trazado de una gráfica circular implica cortar el círculo en las proporciones apropiadas que representen frecuencias relativas. Por ejemplo, la categoría de motos acuáticas es el 46% del total, por lo que el segmento que representa las motos acuáticas debe ser 46% del total (con un ángulo central de $0.46 \times 360° = 166°$).

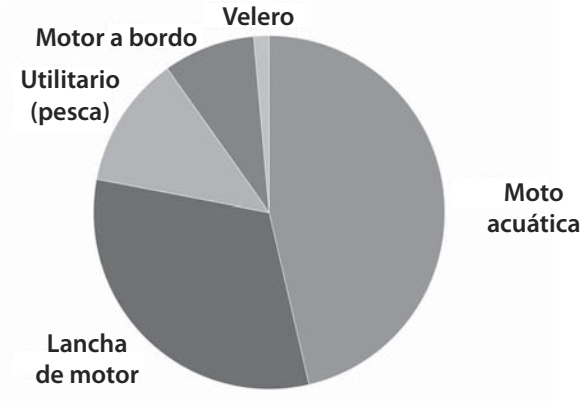


FIGURA 2-9 Gráfica circular de vehículos acuáticos robados

SU TURNO Resuelva el ejercicio 13 "Rectificaciones en revistas".

La gráfica de Pareto en la figura 2-8 y la gráfica circular de la figura 2-9 muestran los mismos datos de diferentes formas, pero la primera muestra de mejor manera los tamaños relativos de los diferentes componentes. El experto en gráficos Edwin Tufte hace la siguiente sugerencia:

> **Nunca utilice gráficas circulares porque desperdician tinta en componentes que no son datos y carecen de una escala apropiada.**

Polígono de frecuencias

Un **polígono de frecuencias** utiliza segmentos de línea conectados a puntos situados directamente encima de los valores de los puntos medios de clase. Un polígono de frecuencias es muy similar a un histograma, pero el polígono de frecuencias utiliza segmentos de línea en vez de barras.

Una variación del polígono de frecuencias básico es el **polígono de frecuencias relativas**, que utiliza frecuencias relativas (proporciones o porcentajes) en la escala vertical. Una ventaja de los polígonos de frecuencias relativas es que dos o más de ellos se pueden combinar en una sola gráfica para facilitar la comparación, como en la figura 2-11 (página 62).

EJEMPLO 6 Polígono de frecuencias de los tiempos de servicio para el almuerzo en McDonald's

Vea en la figura 2-10 el polígono de frecuencias correspondiente a los tiempos de servicio para el almuerzo en McDonald's, que se resumen en la distribución de frecuencias de la tabla 2-2 en la página 42. Las alturas de los puntos corresponden a las frecuencias de clase, los segmentos de línea se extienden hacia la derecha, y la gráfica comienza y termina sobre el eje horizontal.

FIGURA 2-10 Polígono de frecuencias de los tiempos de servicio para el almuerzo en McDonald's

SU TURNO Resuelva el ejercicio 15 "Old Faithful".

EJEMPLO 7 Polígono de frecuencias relativas: Tiempos de servicio para el almuerzo en McDonald's

La figura 2-11 muestra los polígonos de frecuencias relativas de los tiempos de servicio para el almuerzo en auto en McDonald's y Dunkin' Donuts. Se observa que los tiempos de servicio de Dunkin' Donuts son generalmente más bajos (más alejados hacia la izquierda en la gráfica) que los de McDonald's. Esto era de esperarse, dada la distinta naturaleza de sus menús.

continúa

En cifras

2,295,882,327: Cantidad de usuarios de Internet en el mundo.

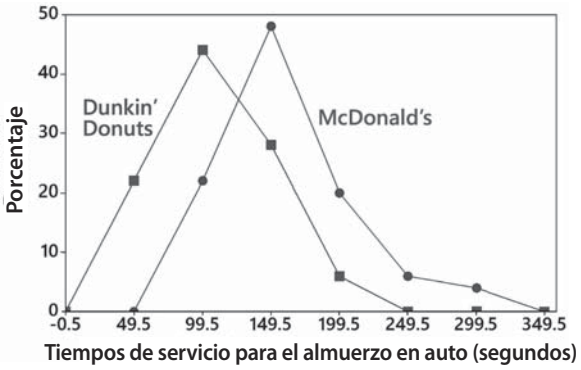


FIGURA 2-11 Polígonos de frecuencias relativas para McDonald's y Dunkin' Donuts

Gráficas que engañan

Por lo general, las gráficas engañosas se usan para mentirle a la gente, y realmente no deseamos que los estudiantes de estadística estén entre las personas susceptibles a tales engaños. Las gráficas se deben trazar de una manera justa y objetiva. Se debe dejar que los lectores hagan sus propios juicios, en vez de manipularlos mediante gráficas engañosas. A continuación presentamos dos de las maneras en que las gráficas suelen usarse para representar los datos de manera engañosa.

Eje vertical sin cero

Una gráfica engañosa común implica el uso de una escala vertical que comienza en algún valor mayor que cero para exagerar las diferencias entre los grupos.

EJE SIN CERO: Siempre examine cuidadosamente una gráfica para ver si el eje vertical empieza en algún punto distinto de cero, de modo que las diferencias sean exageradas.

EJEMPLO 8 Eje sin cero

Las figuras 2-12(a) y 2-12(b) se basan en los mismos datos de un ensayo clínico con OxyContin (oxycodone), un medicamento usado para tratar el dolor de moderado a severo. Los resultados de ese ensayo clínico incluyeron el porcentaje de sujetos que experimentaron náuseas en un grupo de tratamiento con OxyContin y otro grupo al que se administró un placebo.

Si se utiliza una escala vertical que comienza en 10% en lugar de 0%, como en la figura 2-12(a), se exagera groseramente la diferencia entre los dos grupos. La figura 2-12(a) hace parecer que aquellos que usan OxyContin experimentan náuseas a una tasa que es aproximadamente 12 veces mayor que la tasa de los sujetos que toman un placebo, pero la figura 2-12(b) muestra que la relación verdadera es de aproximadamente 2:1, no 12:1. Tal vez alguien quiere desalentar el uso recreativo de OxyContin influyendo en las personas para que piensen que el problema con las náuseas es mucho mayor de lo que realmente es. El objetivo puede ser honesto, pero el uso de una gráfica engañosa no es la forma correcta de lograr ese objetivo.

SU TURNO Resuelva el ejercicio 17 "Vehículos auto-conducidos"

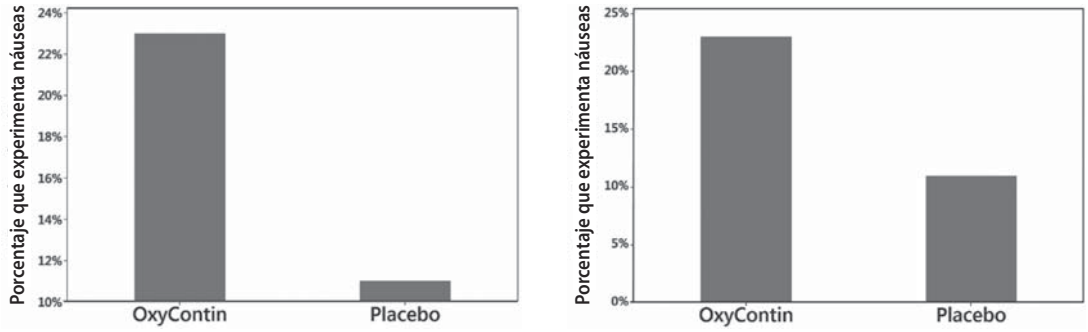


FIGURA 2-12 Náuseas en un ensayo clínico

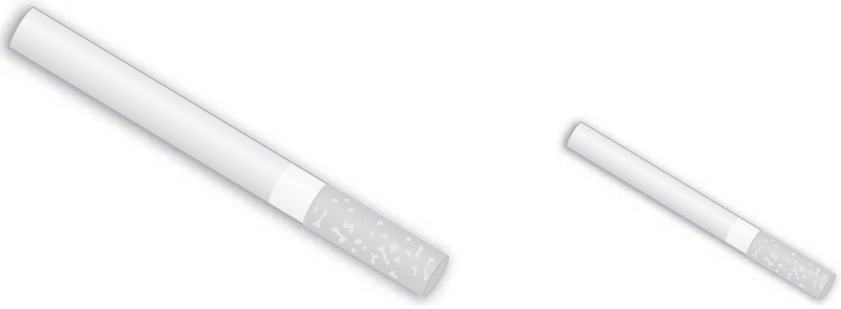
Pictogramas

El trazado de *pictogramas* suele ser engañoso. Los datos que son de naturaleza unidimensional (como las cantidades presupuestarias) se representan a menudo con objetos bidimensionales (como billetes de dólar) u objetos tridimensionales (como pilas de monedas, casas o barriles). Mediante el uso de pictogramas, los dibujantes pueden crear falsas impresiones que distorsionan groseramente las diferencias mediante el uso de sencillos principios de geometría básica: (1) al duplicar cada lado de un cuadrado, su área no sólo se duplica; aumenta por un factor de *cuatro*; (2) cuando se duplica cada lado de un cubo, su volumen no sólo se duplica, sino que aumenta en un factor de *ocho*.

PICTOGRAMAS: Al examinar los datos representados con un pictograma, determine si el gráfico es engañoso porque los objetos de área o volumen se usan para representar cantidades que son realmente unidimensionales. (Los histogramas y las gráficas de barras representan datos unidimensionales con barras bidimensionales, pero usan barras con el mismo ancho para que el gráfico no sea engañoso.)

EJEMPLO 9 Pictograma de fumadores

Observe la figura 2-13 y note que el cigarrillo más grande es aproximadamente dos veces más largo, dos veces más alto y dos veces más grueso que el cigarrillo más pequeño, por lo que el volumen del cigarrillo mayor es aproximadamente *ocho veces* el volumen del menor. (Los datos provienen de los Centros para el Control y la Prevención de Enfermedades). El cigarrillo mayor *parece* ser ocho veces más grande que el cigarrillo menor, pero los porcentajes reales muestran que la tasa del 37% de fumadores en 1970 es aproximadamente *dos veces* el 18% de 2013.



1970: 37% de los estadounidenses fumaban.

2013: 18% de los estadounidenses fumaban.

FIGURA 2-13 Tabaquismo en adultos estadounidenses

SU TURNO Resuelva el ejercicio 19 “Costo de dar a luz”.

Pensamientos conclusivos

Además de las gráficas que se han analizado en esta sección, hay muchas otras gráficas útiles, algunas de las cuales aún no han sido creadas. El mundo necesita desesperadamente más gente que pueda crear gráficas originales que nos aclaren la naturaleza de los datos. En *The Visual Display of Quantitative Information*, Edward Tufte ofrece los siguientes principios:

- Para conjuntos de datos pequeños de 20 valores o menos, utilice una tabla en vez de una gráfica.
- Una gráfica de datos debe hacer que el lector se concentre en la verdadera naturaleza de los datos, no en otros elementos, como características de diseño llamativas pero distractoras.
- No distorsione los datos; elabore la gráfica para revelar la verdadera naturaleza de los datos.
- La mayor parte de la tinta en una gráfica debe utilizarse para los datos, no para otros elementos de diseño.

CENTRO DE TECNOLOGÍA

Capacidades gráficas

Acceda a los complementos de software, videos y conjuntos de datos en **www.pearsonenespañol.com/triola**

En lugar de enumerar instrucciones para cada tipo de gráfica, las siguientes listas identifican las gráficas que se pueden generar mediante el uso de las diferentes tecnologías.

Statdisk

- Histogramas
- Gráficas circulares
- Diagramas de dispersión

Minitab

- Histogramas
- Gráficas de puntos
- Diagramas de tallo y hojas
- Gráficas de series de tiempo
- Gráficas de barras
- Gráficas de Pareto
- Gráficas circulares
- Polígonos de frecuencias
- Diagramas de dispersión

StatCrunch

- Histogramas
- Gráficas de puntos
- Diagramas de tallo y hojas
- Gráficas de barras
- Gráficas circulares
- Diagramas de dispersión

Calculadora TI-83/84 Plus

- Histogramas
- Gráficas de series de tiempo
- Polígonos de frecuencias
- Diagramas de dispersión

Excel

- Histogramas
- Gráficas de series de tiempo
- Gráficas de barras
- Gráficas de Pareto
- Gráficas circulares
- Diagramas de dispersión

2-3 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Temperaturas corporales A continuación se listan temperaturas corporales (°F) de adultos sanos. ¿Por qué una gráfica de estos datos no sería muy efectiva para ayudarnos a entender los datos?

98.6 98.6 98.0 98.0 99.0 98.4 98.4 98.4 98.4 98.6

2. Datos de respuesta voluntaria Si tenemos una gran muestra de respuesta voluntaria consistente en los pesos de sujetos que optaron por responder a una encuesta publicada en Internet, ¿puede una gráfica ayudar a superar la deficiencia de que la muestra sea de respuesta voluntaria?

3. Ética Hay datos que muestran que fumar es perjudicial para la salud. Dado que se podría ayudar a las personas y salvar vidas mediante la reducción del tabaquismo, ¿es ético graficar los datos de una manera que sea engañosa al exagerar los riesgos para la salud del tabaquismo?

4. CVDOT La sección 2-1 introdujo características importantes de los datos resumidas por el acrónimo CVDOT. ¿Qué características representan esas letras y cuál gráfica proporciona el mejor discernimiento de la última de esas características?

Gráficas de puntos. ***En los ejercicios 5 y 6, elabore la gráfica de puntos.***

5. Pulsos A continuación se listan los pulsos (pulsaciones por minuto) de mujeres seleccionadas del conjunto de datos 1 "Datos corporales" en el apéndice B. Todas estos pulsos son números pares. ¿Hay algún pulso que parezca ser un valor atípico? ¿Cuál es ese valor?

80 94 58 66 56 82 78 86 88 56 36 66 84 76 78 64 66 78 60 64

6. Presión arterial diastólica A continuación se listan las mediciones de presión arterial diastólica (mm Hg) de mujeres seleccionadas del conjunto de datos 1 "Datos corporales" en el apéndice B. Todos los valores son números pares. ¿Existen valores atípicos? Si es así, identifique sus valores.

62 70 72 88 70 66 68 70 82 74 90 62
70 76 90 86 60 78 82 78 84 76 60 64

Diagramas de tallo y hojas. ***En los ejercicios 7 y 8, elabore el diagrama de tallo y hojas.***

7. Pulsos Considere los datos listados en el ejercicio 5. ¿Cómo se ordenan los datos en el diagrama de tallo y hojas?

8. Presión arterial diastólica Considere los datos listados en el ejercicio 6. Identifique los dos valores que están más cerca del centro si los datos se ordenan de menor a mayor. (Estos valores se usan a menudo para encontrar la *mediana*, que se define en la sección 3-1).

Gráficas de series de tiempo. ***En los ejercicios 9 y 10, elabore la gráfica de series de tiempo.***

9. Brecha de género en el salario A continuación se listan las medianas de las ganancias de las mujeres como porcentaje de las ganancias medianas de los hombres en los últimos años a partir de 1990. ¿Hay una tendencia? ¿Cómo parece afectar a las mujeres?

71.6 69.9 70.8 71.5 72.0 71.4 73.8 74.2 73.2 72.3 73.7
76.3 76.6 75.5 76.6 77.0 76.9 77.8 77.1 77.0 77.4 77.0

10. Jonrones A continuación se listan las cantidades de jonrones conectados en las Ligas Mayores de Béisbol, cada año, a partir de 1990 (listados en orden por filas). ¿Hay una tendencia?

3317 3383 3038 4030 3306 4081 4962 4640 5064 5528 5693 5458
5059 5207 5451 5017 5386 4957 4878 5042 4613 4552 4934 4661

Gráficas de Pareto. ***En los ejercicios 11 y 12, elabore la gráfica de Pareto.***

11. Rectificaciones en revistas En un estudio de las rectificaciones en revistas biomédicas, 436 fueron por error, 201 por plagio, 888 por fraude, 291 por publicaciones duplicadas y 287 por otras causas (con base en los datos de "Recuento de mala praxis para la mayoría de rectificaciones de revistas científicas" (*Misconducts Accounts for the Majority of Retracted Scientific Publications*", de Fang, Steen, Casadevall, *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, núm. 3). Entre tales rectificaciones, ¿la mala conducta (fraude, duplicación, plagio) parece ser un factor importante?

12. Obtención de un empleo En una encuesta, a los sujetos que buscaban un empleo se les preguntó a quién debían enviar una nota de agradecimiento después de tener una entrevista de trabajo. Los resultados fueron los siguientes: 210 dijeron que sólo a la persona con la que pasaron la mayor parte del tiempo, 396 dijeron que a todos los que conocieron, 40 dijeron que sólo a la persona de más alto nivel, 15 dijeron que a la persona con la que tuvieron la mejor conversación y 10 dijeron que no envían notas de agradecimiento (basado en datos de TheLadders.com). Comente los resultados.

Gráficas circulares. ***En los ejercicios 13 y 14, elabore la gráfica circular.***

13. Rectificaciones en revistas Utilice los datos del ejercicio 11 "Rectificaciones en revistas".

14. Obtención de un empleo Utilice los datos del ejercicio 12 "Obtención de un empleo".

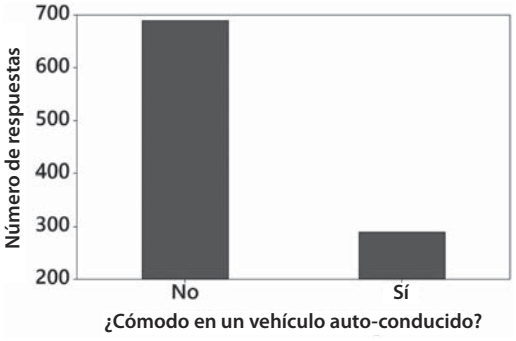
Polígono de frecuencias. ***En los ejercicios 15 y 16, elabore los polígonos de frecuencias.***

15. Old Faithful Utilice la distribución de frecuencias del ejercicio 11 en la sección 2-1 de la página 49 para trazar un polígono de frecuencias. ¿La gráfica sugiere que la distribución es asimétrica? ¿De qué manera lo es?

16. Tornados Utilice la distribución de frecuencias del ejercicio 12 en la sección 2-1 de la página 49 para trazar un polígono de frecuencias. ¿La gráfica sugiere que la distribución es asimétrica? ¿De qué manera lo es?

Gráficas engañosas. ***En los ejercicios 17 a 20, identifique por qué la gráfica es engañosa.***

17. Vehículos auto-conducidos En una encuesta a adultos, se preguntó a los sujetos si se sentían cómodos en un vehículo auto-conducido. La gráfica adjunta muestra los resultados (con base en datos de TE Connectivity).



18. Tarifa del metro En 1986, la tarifa del metro de la Ciudad de Nueva York era de $1, y el costo actual es de $2.50, así que el precio de 1986 se ha multiplicado por 2.5. En la gráfica adjunta, el billete grande es 2.5 veces más alto y 2.5 veces más ancho que el billete pequeño.

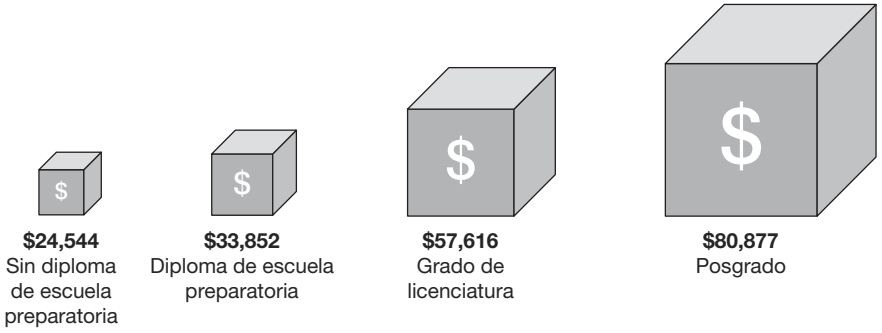
Tarifa del metro en 1986

Tarifa actual del metro

19. Costo de dar a luz De acuerdo con la Agencia para la Investigación de la Salud y el Proyecto de Costo y Utilización de la Salud de Calidad, el costo típico de un parto por cesárea es de $4500 y el de un parto vaginal es de $2600. Observe la siguiente ilustración.

Costo del parto por cesárea: $4500 **Costo del parto vaginal: $2600**

20. Ingresos y grados académicos La gráfica adjunta presenta trabajadores con varios grados académicos, junto con sus niveles de ingreso.



2-3 Más allá de lo básico

21. Diagramas de tallo y hojas expandidas Un diagrama de tallo y hojas se puede *condensar* combinando las filas adyacentes. Podríamos usar un tallo de "6-7" en vez de tallos separados de 6 y 7. Cada fila en el diagrama de tallo y hojas condensado debe incluir un asterisco para separar los dígitos asociados con los diferentes valores del tallo. Un diagrama de tallo y hojas se puede *expandir* subdividiendo las filas en aquellas con hojas que tienen dígitos de 0 a 4 y las que tienen hojas con dígitos de 5 a 9. Si se consideran las temperaturas corporales a partir de las 12 AM del día 2 listadas en el conjunto de datos 3 "Temperaturas corporales" en el apéndice B, se observa que las tres primeras filas de un diagrama de tallo y hojas expandido tienen tallos de 96 (para hojas entre 5 y 9 inclusive), 97 (para hojas entre 0 y 4 inclusive) y 97 (para hojas entre 5 y 9 inclusive). Elabore el diagrama completo de tallo y hojas expandido para las temperaturas corporales a partir de las 12 AM del día 2, que se listan en el conjunto de datos 3 "Temperaturas corporales" en el apéndice B.

2-4 Diagramas de dispersión, correlación y regresión

Concepto clave En esta sección se presenta el análisis de datos muestrales *pareados*. En la parte 1 de esta sección se analiza la *correlación* y el papel de una gráfica llamada *gráfica de dispersión*. En la parte 2 se proporciona una introducción al uso del *coeficiente de correlación lineal*. En la parte 3 se hace un breve análisis de la *regresión lineal*, que implica la ecuación y la gráfica de la recta que mejor se ajusta a los datos muestrales emparejados.

Todos los principios estudiados en esta sección se analizan con mayor detalle en el capítulo 10, pero esta sección sirve como una introducción rápida a algunos conceptos impor-

tantes de la correlación y la regresión. La presente sección no incluye detalles para ejecutar cálculos manuales, los cuales se realizan con poca frecuencia. Las instrucciones para utilizar la tecnología y obtener los resultados deseados se incluyen en el capítulo 10.

PARTE 1 Diagrama de dispersión y correlación

Nuestro objetivo en esta sección es explorar si existe una *correlación*, o asociación, entre dos variables. Comenzamos con definiciones básicas.

DEFINICIONES

Existe una **correlación** entre dos variables cuando los valores de una variable están de alguna manera asociados con los valores de la otra.

Existe una **correlación lineal** entre dos variables cuando hay una correlación y los puntos graficados de los datos pareados dan como resultado un patrón que puede aproximarse mediante una línea recta.

Un **diagrama de dispersión** (o **gráfica de dispersión**) es un diagrama de datos cuantitativos pareados (x, y) con un eje x horizontal y un eje vertical y; el eje horizontal se utiliza para la primera variable (x) y el eje vertical se usa para la segunda variable (y).

PRECAUCIÓN La presencia de una correlación entre dos variables no es evidencia de que una de las variables *cause* la otra. Podríamos encontrar una correlación entre el consumo de cerveza y el peso, pero no podemos concluir a partir de la evidencia estadística que beber cerveza tiene un efecto directo sobre el peso.

¡La correlación no implica causalidad!

Un diagrama de dispersión puede ser muy útil para determinar si existe una correlación (o relación) entre dos variables. (Este tema se analiza de manera extensa durante el estudio de la correlación en la sección 10-1).

EJEMPLO 1 Correlación: Circunferencia de la cintura y del brazo

El conjunto de datos 1 "Datos corporales" en el apéndice B incluye circunferencias de la cintura (cm) y del brazo (cm) de sujetos adultos seleccionados al azar. La figura 2-14 es un diagrama de dispersión de las mediciones de cintura y brazo pareadas. Los puntos muestran un patrón de valores crecientes de izquierda a derecha. Este patrón sugiere que existe una correlación o relación entre las circunferencias de la cintura y las del brazo.

SU TURNO Resuelva el ejercicio 7 "Peso del automóvil y consumo de combustible".

EJEMPLO 2 Sin correlación: Peso y pulso

El conjunto de datos 1 "Datos corporales" en el apéndice B incluye pesos (kg) y pulsaciones (latidos por minuto) de sujetos adultos seleccionados al azar. La figura 2-15 es un diagrama de dispersión de las mediciones de peso y pulso emparejadas. Los puntos en la figura 2-15 no muestran ningún patrón obvio, y esta carencia de un patrón sugiere que no hay correlación o relación entre los pesos y los pulsos.

SU TURNO Resuelva el ejercicio 8 "Estaturas de padres e hijos".

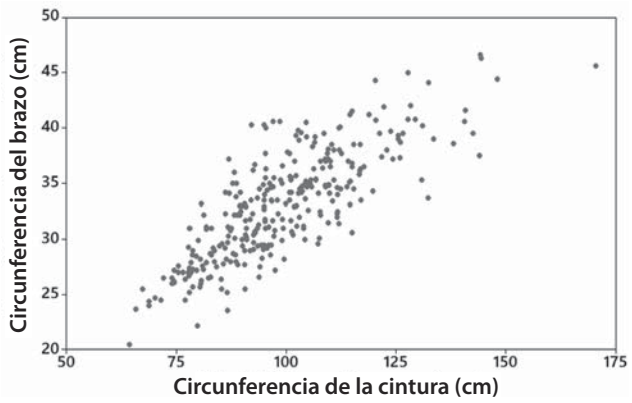


FIGURA 2-14 Circunferencias de cintura y brazo
Correlación: El patrón distintivo de los puntos graficados sugiere que hay una correlación entre las circunferencias de la cintura y las del brazo.

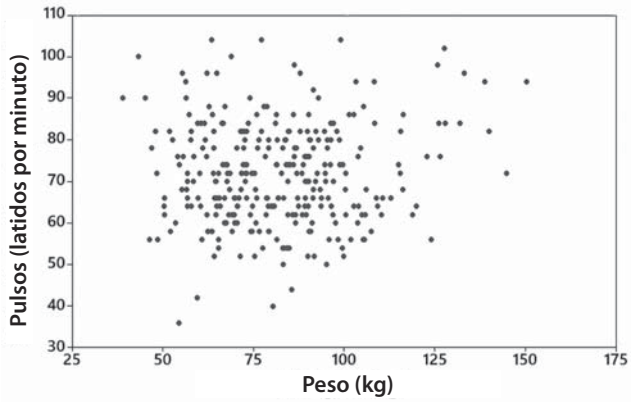


FIGURA 2-15 Pesos y pulsos
Sin correlación: Los puntos trazados no muestran un patrón distintivo, por lo que parece que no hay correlación entre los pesos y los pulsos.

EJEMPLO 3 Conglomerados y un espacio

Considere el diagrama de dispersión de la figura 2-16. Representa los datos apareados que consisten en el peso (gramos) y el año de fabricación de cada uno de 72 centavos. Este diagrama de dispersión muestra dos grupos muy distintos separados por un espacio, lo que puede explicarse por la inclusión de dos poblaciones diferentes: Las monedas de antes de 1983 son 97% de cobre y 3% de zinc, pero las monedas posteriores a 1983 son 2.5% de cobre y 97.5% de zinc. Si se ignora la característica de los conglomerados, podríamos pensar incorrectamente que hay una relación entre el peso de un centavo y el año en que se hizo. Si examinamos los dos grupos por separado, vemos que *no* parece haber una relación entre los pesos de los centavos y los años en que se produjeron.

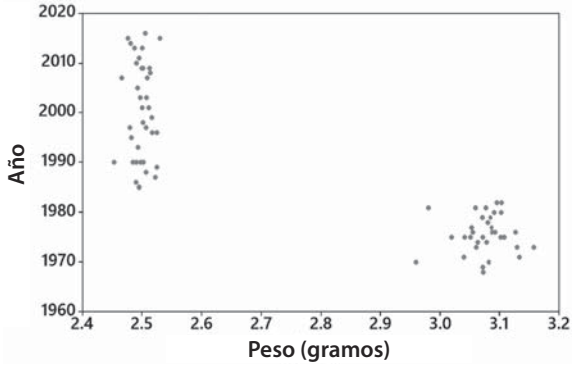


FIGURA 2-16 Pesos de centavos y años de producción

Los tres ejemplos anteriores implican tomar decisiones sobre una correlación basada en juicios subjetivos de diagramas de dispersión, pero la parte 2 presenta el coeficiente de *correlación lineal* como una medida que puede ayudarnos a tomar tales decisiones de manera más objetiva. Mediante el uso de datos pareados, es posible calcular el valor del *coeficiente de correlación lineal r*.

PARTE 2 Coeficiente de correlación lineal *r*

DEFINICIÓN

El **coeficiente de correlación lineal** se expresa con *r*, y mide la fuerza de la asociación lineal entre dos variables.

El valor de un coeficiente de correlación lineal *r* se puede calcular manualmente aplicando la fórmula 10-1 o la fórmula 10-2 que se encuentran en la sección 10-l de la página 473, pero en la práctica, *r* se encuentra casi siempre utilizando software estadístico o una calculadora adecuada.

Utilización de *r* para determinar la correlación

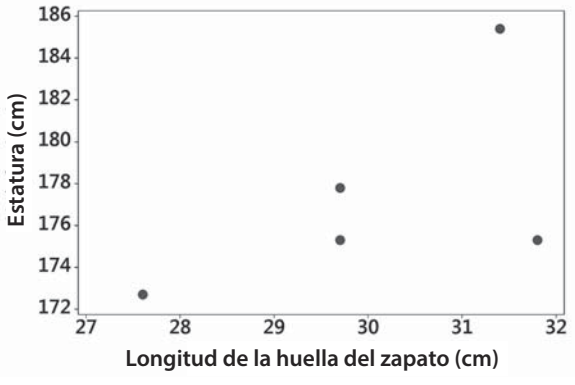
El valor calculado del coeficiente de correlación lineal está siempre entre -1 y 1. Si *r* es cercano a -1 o cercano a 1, parece haber una correlación, pero si *r* es cercano a 0, no parece haber una correlación lineal. Para los datos representados en el diagrama de dispersión de la figura 2-14, $r = 0.802$ (algo cercano a 1), y los datos en el diagrama de dispersión de la figura 2-15 resultan en $r = 0.082$ (muy cercano a 0). Estas descripciones de "cercano a" -1 o 1 o 0 son vagas, pero hay otros criterios objetivos. Por ahora se utilizará una tabla de valores especiales (tabla 2-11) para decidir si existe una correlación lineal. Vea el siguiente ejemplo que ilustra la interpretación del coeficiente de correlación lineal *r*.

EJEMPLO 4 ¿Correlación entre las longitudes de las huellas del zapato y las estaturas?

Considere los datos de la tabla 2-10 (con datos del conjunto de datos 2 "Pie y estatura" del apéndice B). A partir del diagrama de dispersión adjunto de datos pareados en la tabla 2-10, no está muy claro si existe una correlación lineal. La pantalla de resultados de Statdisk muestra que el coeficiente de correlación lineal tiene el valor de $r = 0.591$ (redondeado).

TABLA 2-10 Longitudes de la huella del zapato y estatura de hombres

Longitud de la huella del zapato (cm)	29.7	29.7	31.4	31.8	27.6
Estatura (cm)	175.3	177.8	185.4	175.3	172.7



Statdisk

Sample size, n: 5
Degrees of freedom: 3

Correlation Results:
Correlation coeff, r: 0.5912691
Critical r: ±0.8783393
P-value (two-tailed): 0.29369

Regression Results:
Y= b0 + b1x:
Y Intercept, b0: 125.4073
Slope, b1: 1.727452

Total Variation: 95.02
Explained Variation: 33.21891
Unexplained Variation: 61.80109
Standard Error: 4.538762
Coeff of Det, R^2: 0.3495991

SU TURNO Resuelva el ejercicio 9 "Coeficiente de correlación lineal".

En el ejemplo 4, se sabe por la pantalla de Statdisk que al utilizar los cinco pares de datos de la tabla 2-10, el coeficiente de correlación lineal se calcula como $r = 0.591$. Utilice los siguientes criterios para interpretar dichos valores.

Uso de la tabla 2-11 para interpretar r: Considere los valores críticos de la tabla 2-11 como positivos y negativos y dibuje una gráfica similar a la figura 2-17. Utilice los valores de la tabla para determinar si un valor de un coeficiente de correlación lineal r es "cercano a" 0 o "cercano a" -1 o "cercano a" 1 aplicando los siguientes criterios:

Correlación Si el coeficiente de correlación lineal calculado r está en la región de la cola izquierda o derecha más allá del valor de la tabla para esa cola, concluya que hay evidencia suficiente para apoyar la afirmación de una correlación lineal.

Sin correlación Si el coeficiente de correlación lineal calculado r está entre los dos valores críticos, concluya que no hay evidencia suficiente para apoyar la afirmación de una correlación lineal.

TABLA 2-11 Valores críticos del coeficiente de correlación lineal r

Número de pares de datos n	Valor crítico de r
4	0.950
5	**0.878**
6	0.811
7	0.754
8	0.707
9	0.666
10	0.632
11	0.602
12	0.576

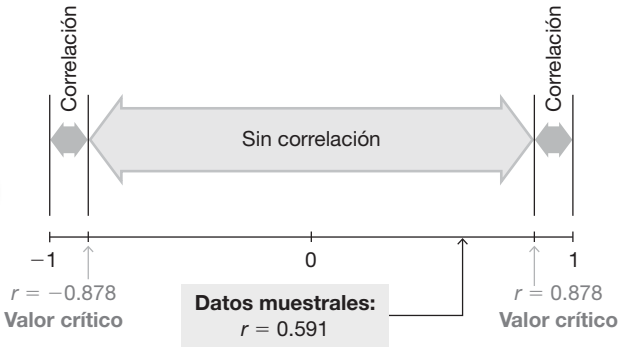


FIGURA 2-17 Valores críticos de la Tabla 2-11 y el valor calculado de *r*

La figura 2-17 muestra que el coeficiente de correlación lineal de $r = 0.591$ calculado a partir de los datos de muestrales pareados es un valor que se encuentra entre los valores críticos de $r = -0.878$ y $r = 0.878$ (se encuentran en la tabla 2-11). La figura 2-17 muestra que podemos considerar que el valor de $r = 0.591$ es cercano a 0 en lugar de estar cerca de -1 o cerca de 1. Por lo tanto, no tenemos evidencia suficiente para concluir que hay una correlación lineal entre las longitudes de la huella del zapato y las estaturas de los hombres.

Valores de *P* para determinar la correlación lineal

En el ejemplo 4, se utilizó el valor calculado del coeficiente de correlación lineal $r = 0.591$ y se comparó con los valores críticos de r de ± 0.878 encontrados en la tabla 2-11. (Vea la figura 2-17). En el mundo real de las aplicaciones estadísticas, el uso de estas tablas es casi obsoleto. La sección 10-1 describe un enfoque más común que se basa en "valores de P" en lugar de tablas. La pantalla de Statdisk que acompaña al ejemplo 4 muestra que el valor de P es 0.29369, o 0.294, una vez que se redondea. Los valores de P se introducen en el capítulo 8, pero a continuación se da una definición preliminar adecuada para el contexto de esta sección:

DEFINICIÓN

Si no hay correlación lineal entre dos variables, el ***valor de P*** es la probabilidad de obtener datos muestrales pareados con un coeficiente de correlación lineal r que sea al menos tan extremo como el obtenido a partir de los datos muestrales emparejados.

Muertes de policías en persecuciones automovilísticas

USA Today investigó el informe anual sobre el número de policías muertos durante persecuciones automovilísticas. Se encontró que la Oficina Federal de Investigaciones (FBI) contabilizó 24 muertes en los últimos 35 años, pero otros registros muestran que hubo 371 muertes durante ese período. El reportero de *USA Today* Thomas Frank escribió que "el bajo conteo es uno de los ejemplos más extremos de la incapacidad del gobierno federal para rastrear con precisión las muertes violentas y ha llevado al FBI a minimizar el peligro de que la policía persiga a los automovilistas". Aparentemente, el FBI categorizaba estas muertes como accidentes automovilísticos en lugar de designarlas como muertes de policías ocurridas durante una persecución automovilística.

Con base en el ejemplo 4 y la pantalla de resultados de Statdisk que muestran un valor P de 0.294, sabemos que hay una probabilidad de 0.294 (o una probabilidad de 29.4%) de obtener un coeficiente de correlación lineal $r = 0.591$ o más extremo, suponiendo que no hay correlación lineal entre la longitud de la huella del zapato y la estatura. (Los valores de r que son "más extremos" que 0.591 son los mayores que 0.591 y los menores que -0.591).

Interpretación de un valor *P*. El valor P de 0.294 del ejemplo 4 es alto. Muestra que existe una alta probabilidad de obtener un coeficiente de correlación lineal de $r = 0.591$ (o más extremo) por casualidad cuando no hay correlación lineal entre las dos variables. Debido a que la probabilidad de obtener $r = 0.591$ o un valor más extremo es tan alta (29.4% de probabilidad), concluimos que no hay suficiente evidencia para afirmar que existe una correlación lineal entre las longitudes de la huella del zapato y las estaturas de los hombres.

Sólo un valor de *P pequeño*, como 0.05 o menor (o una probabilidad de 5% o menos), sugiere que *no* es probable que los resultados de la muestra ocurran por casualidad cuando no hay correlación lineal, por lo que un valor de *P* pequeño apoya la conclusión de que existe una correlación lineal entre las dos variables.

EJEMPLO 5 ¿Correlación entre las longitudes de la huella del zapato y las estaturas?

En el ejemplo 4 se usaron sólo cinco pares de datos del conjunto 2 "Pie y estatura" en el apéndice B. Si se utilizan las longitudes de la huella del zapato y las estaturas de los 40 elementos listados en el conjunto de datos 2 del apéndice B, se obtiene el diagrama de dispersión mostrado en la figura 2-18 y los resultados de Minitab mostrados en la pantalla adjunta. El diagrama de dispersión muestra un patrón distintivo en lugar de tener puntos dispersos por todas partes. Además, se observa que el valor del coeficiente de correlación lineal es $r = 0.813$, y el valor de P es 0.000 redondeado a tres decimales. Debido a que el valor P de 0.000 es *pequeño*, se tiene evidencia suficiente para concluir que existe una correlación lineal entre las longitudes de las huellas del zapato y las estaturas.

En el ejemplo 4, con sólo cinco pares de datos, no teníamos suficiente evidencia para concluir que existe tal correlación lineal, pero con 40 pares de datos, aquí existe evidencia suficiente para concluir la existencia de correlación lineal.

Minitab

```
Pearson correlation of Shoe Print Length and Height = 0.813
P-Value = 0.000
```

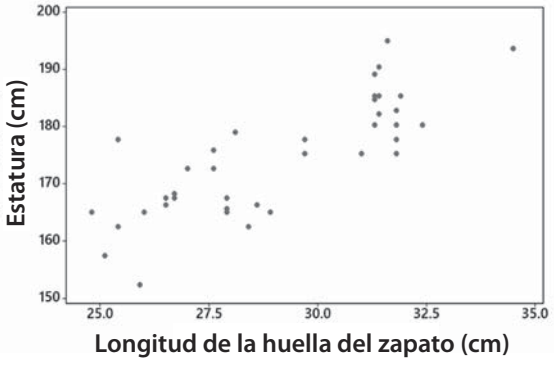


FIGURA 2-18 Diagrama de dispersión de 40 pares de datos

PARTE 3 Regresión

Cuando concluimos que parece haber una correlación lineal entre dos variables (como en el ejemplo 5), podemos encontrar la ecuación de la recta que mejor se ajusta a los datos muestrales, y esa ecuación se puede utilizar para predecir el valor de una variable cuando se da un valor específico de la otra variable. Con base en los resultados del ejemplo 5, podemos predecir la estatura de alguien dada la longitud de la huella de su zapato (que puede haber sido encontrada en una escena del crimen).

En vez de usar el formato de la ecuación lineal $y = mx + b$ que aprendimos en los cursos previos de matemáticas, usamos el formato siguiente.

DEFINICIÓN

Dada una colección de datos muestrales pareados, la **línea de regresión** (o *línea de mejor ajuste* o *línea de mínimos cuadrados*) es la recta que "mejor" se ajusta a la dispersión de los datos. (El criterio específico para la "mejor" línea recta de ajuste es la propiedad de "mínimos cuadrados" descrita en la sección 10-2).

La **ecuación de regresión**

$$\hat{y} = b_0 + b_1 x$$

describe algebraicamente la línea de regresión.

La sección 10-2 proporciona una buena razón para usar el formato $\hat{y} = b_0 + b_1 x$ en vez del formato $y = mx + b$. La sección 10-2 también proporciona fórmulas que podrían utilizarse para identificar los valores de la intersección en y, b_0, y de la pendiente, b_1, pero los valores se encuentran normalmente mediante el uso de un software estadístico o una calculadora adecuada.

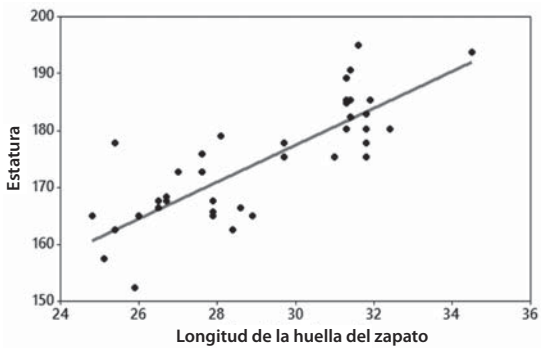
EJEMPLO 6 Línea de regresión

El ejemplo 5 incluyó un diagrama de dispersión de los 40 pares de longitudes de huella del zapato y estaturas del conjunto de datos 2 "Pie y estatura" del apéndice B. La figura 2-19 de la página siguiente muestra ese mismo diagrama de dispersión, pero incluye la gráfica de la línea de regresión. También se muestra la pantalla de Statdisk para los 40 pares de datos.

A partir de la pantalla de Statdisk, se observa que la forma general de la ecuación de regresión tiene un intersección y de $b_0 = 80.9$ (redondeada) y una pendiente $b_1 = 3.22$ (redondeada), por lo que la ecuación de la línea de regresión mostrada en la figura 2-19 es $\hat{y} = 80.9 + 3.22x$. Podría ser útil expresar esta ecuación más claramente usando los nombres de las variables:

$$\text{Estatura} = 80.9 + 3.22\ (\text{Longitud de la huella del zapato})$$

Observe que la *ecuación* muestra la intersección y de 80.9 que no aparece en la escala vertical de la *gráfica*. La escala vertical más a la izquierda de la figura 2-19 no es el eje real de y que pasa por 0 en el eje x. Si la gráfica se extendiera a la izquierda, la línea de regresión intersecaría al eje y real a la altura de $y = 80.9$ cm.



Statdisk

```
Correlation Results:
Correlation coeff, r: 0.812948
Critical r: ±0.3120061
P-value (two-tailed): 0.000

Regression Results:
Y= b0 + b1x:
Y Intercept, b0: 80.93041
Slope, b1: 3.218561
```

FIGURA 2-19 Línea de regresión

2-4 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Correlación lineal En esta sección usamos r para expresar el valor del coeficiente de correlación lineal. ¿Por qué nos referimos a este coeficiente de correlación como *lineal*?

2. Causalidad Un estudio ha demostrado que existe una correlación entre el peso corporal y la presión arterial. Los pesos corporales más altos están asociados con los niveles más altos de presión arterial. ¿Podemos concluir que el aumento de peso es una causa del aumento de la presión arterial?

3. Diagrama de dispersión ¿Qué es un diagrama de dispersión y cómo nos ayuda?

4. Estimación de *r* En cada uno de los siguientes casos, estime el valor del coeficiente de correlación lineal r para los datos pareados que se dan y que corresponden a 50 adultos seleccionados al azar.

a. Sus estaturas se miden en pulgadas (x) y esas mismas estaturas se registran en centímetros (y).

b. Se miden sus puntuaciones de IQ (x) y sus estaturas en centímetros (y).

c. Se miden sus pulsos (x) y sus puntuaciones de IQ (y).

d. Se miden sus estaturas en centímetros (x) y se listan de nuevo las mismas estaturas, pero con un signo negativo precediendo cada dato (y).

Diagrama de dispersión ***En los Ejercicios 5 a 8, use los datos muestrales para trazar un diagrama de dispersión. Utilice la primera variable para el eje x. Con base en el diagrama de dispersión, ¿qué conclusión obtiene sobre una correlación lineal?***

5. Volumen cerebral e IQ La tabla lista los volúmenes del cerebro (cm^3) y las puntuaciones de IQ para cinco varones (del conjunto de datos 8 "IQ y tamaño del cerebro" del apéndice B).

Volumen del cerebro (cm^3)	1173	1067	1347	1029	1204
IQ	101	93	94	97	113

6. Medidas de un oso La tabla lista los tamaños de tórax (distancia alrededor del pecho en pulgadas) y pesos (libras) que se midieron en osos anestesiados (conjunto de datos 9 "Mediciones de un oso" en el apéndice B).

Pecho (pulg.)	26	45	54	49	35	41	41
Peso (lb)	80	344	416	348	166	220	262

7. Peso del auto y consumo de combustible La tabla lista los pesos (en libras) y los consumos de combustible en carretera (mpg) para los autos Hyundai Elantra, Nissan Altima, VW Passat, Buick Lucerne, Mercury Grand Marquis, Honda Civic y Honda Accord.

Peso (lb)	2895	3215	3465	4095	4180	2740	3270
Carretera (mpg)	33	31	29	25	24	36	30

8. Estaturas de padres e hijos La tabla lista las alturas de padres y de sus primeros hijos (obtenido de Francis Galton).

Altura del padre (pulg.)	73.0	75.5	75.0	75.0	75.0	74.0	74.0	73.0	73.0	78.5
Altura del primer hijo (pulg.)	74.0	73.5	71.0	70.5	72.0	76.5	74.0	71.0	72.0	73.2

Coeficiente de correlación lineal. ***En los ejercicios 9 a 12, se proporciona el coeficiente de correlación lineal r. Utilice la tabla 2-11 en la página 71 para encontrar los valores críticos de r. Con base en una comparación del coeficiente de correlación lineal r y los valores críticos, ¿qué concluye usted acerca de una correlación lineal?***

9. Considerando los datos del ejercicio 5 "Volumen cerebral e IQ", el coeficiente de correlación lineal es $r = 0.127$.

10. Considerando los datos del ejercicio 6 "Mediciones de un oso", el coeficiente de correlación lineal es $r = 0.980$.

11. Considerando los datos del ejercicio 7 "Peso del auto y consumo de combustible", el coeficiente de correlación lineal es $r = -0.987$.

12. Considerando los datos del ejercicio 8 "Estaturas de padres e hijos", el coeficiente de correlación lineal es $r = -0.017$.

2-4 Más allá de lo básico

Valores de *P*. ***En los ejercicios 13 a 16, escriba un enunciado que interprete el valor P e incluya una conclusión acerca de la correlación lineal.***

13. Considerando los datos del ejercicio 5 "Volumen cerebral e IQ" el valor de *P* es 0.839.

14. Considerando los datos del ejercicio 6 "Mediciones de un oso" el valor de *P* es 0.000.

15. Considerando los datos del ejercicio 7 "Peso del auto y consumo de combustible", el valor de *P* es 0.000.

16. Usando los datos del ejercicio 8 "Estaturas de padres e hijos", el valor de *P* es 0.963.

Examen rápido del capítulo

1. Galletas Revise la distribución de frecuencias adjunta que resume el número de chispas de chocolate en cada galleta de una muestra de Chips Ahoy regulares (del conjunto de datos 28 "Galletas con chispas de chocolate" en el apéndice B). ¿Cuál es la anchura de clase? ¿Es posible identificar los valores de los datos originales?

Chispas de chocolate	Frecuencia
18-20	6
21-23	11
24-26	18
27-29	4
30-32	1

2. Galletas Con base en la misma distribución de frecuencias del ejercicio 1, identifique las fronteras y los límites de la primera clase.

3. Galletas Con base en la misma distribución de frecuencias del ejercicio 1, ¿cuántas galletas se incluyen?

4. Galletas Se crea un diagrama de tallo y hojas para las mismas galletas resumidas en el ejercicio 1, y la primera fila de es diagrama es 1|99. Identifique los valores representados por esa fila del diagrama de tallo y hojas.

5. Computadoras Como gerente de control de calidad en Texas Instruments, usted encuentra que los defectos en las calculadoras tienen varias causas, incluyendo maquinaria desgastada, errores humanos, suministros incorrectos y maltratos durante el embalaje. ¿Cuál de las siguientes gráficas sería la mejor para describir las causas de los defectos: histograma, diagrama de dispersión, gráfica de Pareto, gráfica de puntos, gráfica circular?

6. Distribución de la riqueza En los últimos años, ha habido mucha discusión sobre la distribución de la riqueza entre los adultos de Estados Unidos. Si usted planea realizar una investigación original obteniendo de alguna manera el monto de la riqueza de 3000 adultos seleccionados al azar, ¿qué gráfica sería la mejor para ilustrar la *distribución* de la riqueza?

7. Ensayo de salud En una investigación de la relación entre las presiones arteriales sistólica y diastólica de las mujeres adultas, ¿cuál de las siguientes gráficas es más útil: histograma, gráfica circular, diagrama de dispersión, diagrama de tallo y hojas, gráfica de puntos?

8. Lotería En el juego de lotería Play 4 de Florida, cada día se seleccionan aleatoriamente cuatro dígitos entre 0 y 9 inclusive. Normalmente esperamos que cada uno de los 10 dígitos ocurra alrededor de 1/10 de las veces, y un análisis de los resultados del año pasado muestra que así sucedió. Debido a que los resultados son lo que normalmente esperamos, ¿es correcto decir que la distribución de los dígitos seleccionados es una distribución normal?

9. Cinturones de seguridad La compañía Cinturones de Seguridad Beams fabrica ... bueno, usted ya sabe. Cuando se examina una muestra de cinturones de seguridad en relación con su punto de ruptura (medida en kilogramos), se exploran los datos muestrales. Identifique la característica importante de los datos que falta en la siguiente lista: centro, distribución, valores atípicos, características cambiantes en el tiempo.

10. Cinturones de seguridad Se trazará un histograma a partir de los puntos de ruptura medidos (en libras) de los cinturones de seguridad para automóvil probados. Identifique dos características clave de un histograma de esos valores que sugieren que los datos tienen una *distribución normal*.

Ejercicios de repaso

1. Distribución de frecuencias de temperaturas corporales Elabore una distribución de frecuencias de las 20 temperaturas corporales (°F) que se listan a continuación. (Estos datos pertenecen al conjunto de datos 3 "Temperaturas corporales" del apéndice B.) Utilice una anchura de clase de 0.5 °F y un valor inicial de 97.0 °F.

97.1	97.2	97.5	97.6	97.6	97.8	98.0	98.0	98.2	98.2
98.2	98.3	98.4	98.6	98.6	98.7	98.7	98.9	99.1	99.4

2. Histograma de temperaturas corporales Elabore el histograma que corresponde a la distribución de frecuencias del ejercicio 1. Utilice los puntos medios de clase para la escala horizontal. ¿El histograma sugiere que los datos provienen de una población con distribución normal? ¿Por qué sí o por qué no?

3. Gráfica de puntos de temperaturas corporales Elabore una gráfica de puntos de las temperaturas corporales listadas en el ejercicio 1. ¿Cuál es la mejor manera de ilustrar la distribución de los datos: el histograma del ejercicio 2 o la gráfica de puntos?

4. Diagrama de tallo y hojas de temperaturas corporales Elabore un diagrama de tallo y hojas de las temperaturas corporales listadas en el ejercicio 1. ¿Existen valores atípicos?

5. Temperaturas corporales A continuación se listan las temperaturas de nueve varones medidos a las 8 AM y de nuevo a las 12 AM (del conjunto de datos 3 "Temperaturas corporales" en el apéndice B). Elabore un diagrama de dispersión y, con base en él, diga si parece haber una relación entre las temperaturas de las 8 AM y de las 12 AM.

8 AM	98.0	97.0	98.6	97.4	97.4	98.2	98.2	96.6	97.4
12 AM	98.0	97.6	98.8	98.0	98.8	98.8	97.6	98.6	98.6

6. Medio ambiente

a. Después de recopilar las temperaturas medias (promedio) globales en cada uno de los últimos 100 años, queremos trazar la gráfica más apropiada para estos datos. ¿Qué gráfica sería la mejor?

b. Después de recopilar las temperaturas medias (promedio) globales y la cantidad de emisiones de monóxido de carbono en los últimos 100 años, queremos trazar una gráfica para investigar la asociación entre esas dos variables. ¿Qué gráfica es la mejor?

c. Una investigación de las fuentes de monóxido de carbono incluye vehículos de motor, hornos, fuegos, centrales eléctricas que queman carbón y humo de tabaco. Si queremos trazar una gráfica que ilustre la importancia relativa de estas fuentes, ¿qué gráfica es la mejor?

7. Como que es hora de hacer este ejercicio En una encuesta Marista aplicada a adultos, las siguientes son las palabras o frases que los sujetos encuentran más molestas en una conversación (junto con sus frecuencias de respuesta): como que o como algo (127); nada más digo (81); ya sabes (104); como sea (219); obviamente (35). Elabore una gráfica circular. Identifique una desventaja de una gráfica circular.

8. Como sea Utilice los mismos datos del ejercicio 7 para trazar una gráfica de Pareto. ¿Qué gráfica ilustra de mejor manera los datos: la gráfica de Pareto o la gráfica circular?

Ejercicios de repaso acumulado

En los ejercicios 1 a 6, utilice los siguientes datos, que son los tiempos totales de juego en casa (horas) para todos los equipos de Grandes Ligas en un año reciente (con base en los datos de Baseball Prospectus).

236 237 238 239 241 241 242 245 245 245 246 247 247 248 248
249 250 250 250 251 252 252 253 253 258 258 258 260 262 264

1. Distribución de frecuencias Elabore una distribución de frecuencias. Utilice una anchura de clase de 5 horas y un tiempo de inicio de 235 horas.

2. Distribución de frecuencias Para la distribución de frecuencias del ejercicio 1, busque lo siguiente.

a. Límites de la primera clase

b. Fronteras de la primera clase

c. Punto medio de la primera clase

3. Histograma Elabore el histograma correspondiente a la distribución de frecuencias del ejercicio 1. Para los valores en el eje horizontal, utilice los puntos medios de clase. ¿Cuál de las siguientes expresiones describe de mejor manera la distribución: uniforme, normal, asimétrica a la izquierda, asimétrica a la derecha?

4. Gráfica engañosa Suponga que desea crear el histograma para el ejercicio 3 de manera que exagere las diferencias entre los tiempos. Describa cómo se puede modificar el histograma del ejercicio 3 para lograr esa exageración.

5. Diagrama de tallo y hojas Utilice los tiempos totales de juego para crear un diagrama de tallo y hojas. ¿Qué revela este diagrama sobre la distribución de los datos?

6. Tipo de datos

a. Los tiempos de juego listados se redondean al número entero más cercano. Antes de redondear, ¿los tiempos exactos de juego son datos discretos o datos continuos?

b. ¿Para los tiempos listados, los datos son categóricos o cuantitativos?

c. Identifique el nivel de medición de los tiempos listados: nominal, ordinal, de intervalo o de razón.

continúa

d. ¿Cuál de las siguientes opciones describe mejor los datos muestrales: muestra de respuesta voluntaria, muestra aleatoria, muestra de conveniencia, muestra simple?

e. Los tiempos de juego totales listados son de un año reciente, y se dispone de los datos de todos los años hasta 1950. Dado que los tiempos listados son parte de una colección más grande de tiempos, ¿los datos constituyen una muestra o una población?

Proyecto de tecnología

En este capítulo se estableció que los días de las gráficas encantadoras y primitivas dibujadas a mano habían quedado muy atrás, y la tecnología ahora proporciona herramientas poderosas para generar una gran variedad de gráficas. Por lo tanto, este proyecto sirve como una buena preparación para las presentaciones profesionales que inevitablemente se harán en el futuro.

Los conjuntos de datos completos del apéndice B se pueden descargar de www.pearsonenespañol.com/triola. Statdisk ya incluye los conjuntos de datos. Pueden abrirse mediante paquetes de software estadístico, como Statdisk. Minitab, Excel, SPSS y JMP. Utilice uno de estos paquetes para abrir el conjunto de datos 4 "Nacimientos". Utilice los métodos de este capítulo para explorar y comparar el peso al nacer de las niñas y de los niños. (Debido a que las unidades son gramos, en realidad son medidas de masa y no de peso).

- Obtenga una copia impresa de los dos histogramas.
- Describa las naturalezas de las dos distribuciones (uniforme, normal, asimétrica a la izquierda, asimétrica a la derecha), e identifique los posibles valores atípicos.
- Escriba una breve descripción de sus resultados.

Sugerencia: Los géneros están codificados como 1 para masculino y 0 para femenino, así que ordene todas las filas utilizando la columna de género como base para la clasificación. Después se pueden separar las filas de los niños y de las niñas, usando las funciones de copiar/pegar y cortar del software.

DE LOS DATOS A LA DECISIÓN

Tiempos de servicio en auto de los restaurantes de comida rápida: ¿Quién es el mejor?

El conjunto de datos 25 "Comida rápida" en el apéndice B incluye los tiempos de servicio para el almuerzo y la cena en los restaurantes McDonald's, Burger King, Wendy's y Dunkin' Donuts. Varios ejemplos y ejercicios en este capítulo usan algunos de esos tiempos de servicio. Para este proyecto, utilice todos los tiempos del conjunto.

Pensamiento crítico

Utilice los métodos de este capítulo para responder las siguientes preguntas.

1. ¿Cuál de los cuatro restaurantes parece tener el tiempo de servicio en auto más rápido para el almuerzo?

2. ¿Cuál de los cuatro restaurantes parece tener el tiempo de servicio en auto más rápido para la cena?

3. ¿Los tiempos de servicio en auto para el almuerzo parecen ser diferentes de los tiempos de servicio en auto para la cena? Explique.

4. De acuerdo con las opciones de menú disponibles en los diferentes restaurantes, ¿alguno de ellos tiene una ventaja inherente en relación con los tiempos de servicio? Explique.

5. Considerando las diferencias en las opciones de menú, ¿hay un restaurante que parezca ser más eficiente que los demás? Explique.

Actividades en equipo

1. Actividad fuera de la clase El problema del capítulo usa los tiempos de servicio medidos en el conjunto de datos 25 "Comida rápida" del apéndice B. Vaya a uno o más restaurantes de comida rápida y recopile sus propios tiempos de servicio. Compare los resultados con los encontrados en el conjunto de datos 25 del apéndice B.

2. Actividad en clase Utilice un paquete de galletas con chispas de chocolate; cada estudiante debe recibir dos o tres galletas. Proceda a contar el número de chispas de chocolate en cada galleta. No todas las chispas son visibles, por lo que deben utilizarse "pruebas destructivas" a través de un proceso que implica el consumo. Registre el número de chispas de chocolate para cada galleta y combine todos los resultados. Elabore una distribución de frecuencias, un histograma, una gráfica de puntos y un diagrama de tallo y hojas con los resultados. Dado que las galletas se hicieron a través de un proceso de producción en masa, podríamos esperar que el número de chispas de chocolate por galleta no varíe mucho. ¿Indican eso los resultados? Explique.

3. Actividad en clase En la clase, cada estudiante debe registrar dos frecuencias de pulso contando el número de sus latidos cardíacos en 1 minuto. El primer pulso se debe medir mientras el estudiante está sentado; el segundo, mientras está de pie. Utilizando las pulsaciones contadas mientras están sentados, elabore una distribución de frecuencias y un histograma para los pulsos de los varones, y luego otra distribución de frecuencias e histograma para los pulsos de las mujeres. Utilizando las pulsaciones medidas mientras están de pie, elabore una distribución de frecuencias y un histograma para los pulsos de los varones, y luego otra distribución de frecuencias e histograma para los pulsos de las mujeres. Compare los resultados. ¿Los hombres y las mujeres parecen tener pulsos diferentes? ¿Las pulsaciones contadas al estar sentados parecen ser diferentes de los pulsos mientras están de pie? Utilice una gráfica apropiada para determinar si existe una relación entre el pulso sentado y el pulso de pie.

4. Actividad fuera de la clase Busque periódicos y revistas para encontrar un ejemplo de una gráfica que sea engañosa. Describa por qué lo es. Vuelva a dibujar la gráfica para que represente la información correctamente.

5. Actividad fuera de la clase Encuentre la gráfica de Charles Joseph Minard que describe la marcha de Napoleón a Moscú y viceversa, y explique por qué Edward Tufte dice que "puede ser la mejor gráfica jamás dibujada". (Vea *The Visual Display of Quantitative Information* de Edward Tufte, Grapics Press).

6. Actividad fuera de la clase En *The Visual Display of Quantitative Information* de Edward Tufte (Graphics Press), encuentre la gráfica que apareció en *American Education*, y explique por qué Tufte dice que "esta puede ser la peor gráfica jamás impresa". Elabore una gráfica que resulte efectiva para representar los mismos datos.

3

- **3-1** Medidas de tendencia central
- **3-2** Medidas de variación
- **3-3** Medidas de posición relativa y gráficas de caja

DESCRIPCIÓN, EXPLORACIÓN Y COMPARACIÓN DE DATOS

PROBLEMA DEL CAPÍTULO

¿Qué compañía tiene la mejor velocidad (de transferencia) de datos para teléfonos inteligentes en los aeropuertos?

El conjunto de datos 32, "Velocidades de datos en aeropuertos" del apéndice B, lista las velocidades (de transferencia) de datos medidos por RootMetrics en 50 aeropuertos de Estados Unidos considerando las cuatro principales compañías proveedoras del servicio en ese país (Verizon, Sprint, AT&T y T-Mobile). Todas las velocidades se dan en unidades de megabits (o 1 millón de bits) por segundo, expresados como Mbps. Debido a que las velocidades de datos originales listadas en el conjunto de datos 32 incluyen números decimales como 38.5 Mbps, las gráficas de puntos con los datos no modificados serían un poco complicadas y no tan útiles, pero si redondeamos todos los conjuntos de datos originales, obtendremos la gráfica de puntos mostrada en la figura 3-1. (Una observación de la escala horizontal en la figura 3-1 revela que las velocidades de datos originales han sido redondeadas al entero par más cercano por el software utilizado para crear las gráficas de puntos). Si se utilizan las mismas cuatro escalas horizontales y se apilan las cuatro gráficas de puntos, las comparaciones resultan mucho más sencillas.

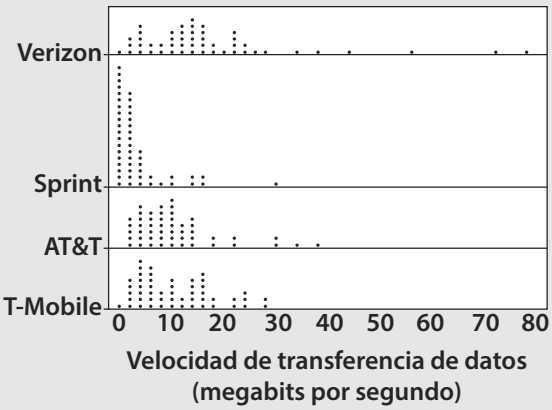


FIGURA 3-1 Gráfica de puntos de las velocidades de datos para Smartphone

El análisis de la figura 3-1 sugiere que Verizon tiene el mejor desempeño global, con velocidades de datos que tienden a ser más altas que las de las otras tres compañías. Pero en vez de confiar únicamente en interpretaciones subjetivas de una gráfica como la de la figura 3-1, este capítulo introduce medidas que son esenciales para cualquier estudio de estadística: la media, la mediana, la desviación estándar y la varianza, que se encuentran entre los datos estadísticos más importantes en el estudio de esta materia. Éstos se usarán para describir, explorar y comparar las velocidades de datos medidos de Verizon, Sprint, AT&T y T-Mobile como se listan en el conjunto de datos 32.

OBJETIVOS DEL CAPÍTULO

Pensamiento crítico e interpretación: más allá de las fórmulas y la aritmética

En este curso moderno de estadística no es tan importante memorizar fórmulas o efectuar cálculos manuales de aritmética. Es posible obtener resultados con una calculadora o software para que podamos concentrarnos en el sentido práctico de los resultados a través del pensamiento crítico. Aunque el presente capítulo incluye pasos detallados para procedimientos importantes, no siempre es necesario dominarlos. No obstante, a menudo es útil llevar a cabo unos cuantos cálculos manuales antes de usar la computadora, con el fin de incentivar la comprensión.

Los métodos y herramientas presentados en este capítulo se llaman con frecuencia métodos de **estadística descriptiva**, puesto que resumen o *describen* las características relevantes de los datos. En los capítulos subsecuentes usamos la **estadística inferencial** para hacer *inferencias* , o generalizaciones, sobre las poblaciones. Los siguientes son los objetivos del capítulo:

3-1 Medidas de tendencia central

- Desarrollar la capacidad de medir el centro de los datos mediante la determinación de la media, la mediana, la moda y la mitad del rango.
- Determinar si un valor atípico tiene un efecto sustancial sobre la media y la mediana.

3-2 Medidas de variación

- Desarrollar la capacidad de medir la variación en un conjunto de datos muestrales mediante la determinación de los valores del rango, la varianza y la desviación estándar.
- Desarrollar la capacidad de interpretar los valores de la desviación estándar aplicando la *regla práctica del rango* para determinar si un valor particular es *significativamente bajo* o *significativamente alto*.

3-3 Medidas de posición relativa y gráficas de caja

- Desarrollar la capacidad de calcular una puntuación *z* y utilizar el resultado para determinar si un valor dado *x* es *significativamente bajo* o *significativamente alto*.
- Identificar valores de los *percentiles* y *cuartiles* de un conjunto de datos.
- Desarrollar la capacidad de construir una gráfica de caja a partir de un conjunto de datos.

3-1 Medidas de tendencia central

Concepto clave El enfoque de esta sección está en obtener un valor que mida el *centro* de un conjunto de datos. En particular, se presentan medidas de tendencia central, incluyendo la *media* y la *mediana*. El objetivo aquí no es sólo encontrar el valor de cada medida de tendencia central, sino también interpretar esos valores. La parte 1 de la presente sección incluye conceptos básicos que deben ser entendidos antes de considerar la parte 2.

PARTE 1 Conceptos básicos de las medidas de tendencia central

En la parte 1 de esta sección introducimos la media, la mediana, la moda y la mitad del rango como diferentes medidas de tendencia central. Tales medidas se utilizan ampliamente para proporcionar valores representativos que "resumen" los conjuntos de datos.

En cifras

$3.19: Monto medio dejado por el ratón de los dientes, con base en una encuesta realizada por Visa. El 10% de los niños con menos suerte no obtiene nada.

DEFINICIÓN

Una **medida de tendencia central** es un valor en medio o en el centro de un conjunto de datos.

Existen diferentes métodos para medir el centro, por lo que tenemos distintas definiciones para ellos. Comenzamos con la media.

Media

Por lo general, la media (o media aritmética) es la más importante de las mediciones numéricas usadas para describir datos, y es lo que la mayoría de las personas llama *promedio*.

DEFINICIÓN

La **media** (o **media aritmética**) de un conjunto de datos es la medida de tendencia central que se encuentra al sumar todos los valores de los datos y dividir el total por el número de datos.

Propiedades importantes de la media

- Las medias muestrales de una misma población tienden a variar menos que otras medidas de tendencia central.
- La media de un conjunto de datos utiliza todos los valores de los datos.
- Una desventaja de la media es que un solo valor extremo (atípico) puede cambiar el valor de la media en forma sustancial. (Con base en la siguiente definición, puede decirse que la media no es *resistente*).

DEFINICIÓN

Un dato estadístico es **resistente** si la presencia de valores extremos (atípicos) no ocasiona que éste cambie mucho.

Cálculo y notación de la media

La definición de la media puede expresarse como en la fórmula 3-1, donde la letra griega Σ (sigma mayúscula) indica que los valores de los datos deben sumarse, por lo que Σx representa la suma de todos los valores de los datos. El símbolo n expresa el **tamaño de la muestra**, que es el número de valores de datos.

FÓRMULA 3-1

$$\text{Media} = \frac{\Sigma x}{n} \quad \begin{matrix} \leftarrow \text{suma de todos los valores de datos} \\ \leftarrow \text{número de valores de datos} \end{matrix}$$

Si los datos son una *muestra* de una población, la media se expresa con $\bar{x}$ (que se pronuncia "x barra"); si los datos son la población entera, la media se expresa mediante μ (letra griega mu minúscula).

NOTACIÓN *Sugerencia:* Los estadísticos muestrales se representan usualmente mediante caracteres latinos, como $\bar{x}$, y los parámetros de población por medio de letras griegas, como μ.

Σ	expresa la *suma* de un conjunto de valores de datos.
x	es la *variable* que generalmente se usa para representar los valores de datos individuales.
n	representa el número de valores de datos en una *muestra*.
N	representa el número de valores de datos en una *población*.
$\bar{x} = \frac{\Sigma x}{n}$	es la media de un conjunto de valores *muestrales*.
$\mu = \frac{\Sigma x}{N}$	es la media de todos los valores en una *población*.

EJEMPLO 1 **Media**

El conjunto de datos 32 "Velocidades de datos en aeropuertos" del apéndice B incluye medidas de las velocidades de datos para teléfonos inteligentes en cuatro compañías. Encuentre la media de los cinco primeros datos de velocidad para Verizon: 38.5, 55.6, 22.4, 14.1 y 23.1 (todos en megabits por segundo, o Mbps).

SOLUCIÓN

La media se calcula mediante la fórmula 3-1. Primero sume los valores de datos, luego divida por el número de valores:

$$\bar{x} = \frac{\Sigma x}{n} = \frac{38.5 + 55.6 + 22.4 + 14.1 + 23.1}{5} = \frac{153.7}{5}$$
$$= 30.74 \text{ Mbps}$$

La media de las cinco primeras velocidades de datos para Verizon es de 30.74 Mbps.

SU TURNO Encuentre la media en el ejercicio 5 "Números de jugadores de fútbol americano".

PRECAUCIÓN Nunca utilice el término *promedio* cuando se refiera a una medida de tendencia central. Esa palabra se utiliza a menudo para la media, pero en ocasiones se usa para otras medidas del centro. Los estadísticos no emplean el término *promedio* y en el resto de este libro no se usará para hacer referencia a una medida de tendencia central específica. El término *promedio* tampoco es utilizado por la comunidad estadística o las revistas profesionales.

Paradoja del tamaño de la clase

Existen al menos dos formas de obtener el tamaño de una clase promedio, y ambas pueden dar resultados muy diferentes. En una universidad, si tomamos la cantidad de estudiantes de 737 clases, obtenemos una media de 40 estudiantes. Sin embargo, si reunimos una lista del tamaño de las clases para cada estudiante y utilizamos esta lista, obtendríamos una media de 147. Esta gran discrepancia se debe al hecho de que existen muchos estudiantes en clases grandes, en tanto que hay pocos estudiantes en clases pequeñas. Sin cambiar el número de clases o de profesores, podríamos reducir el tamaño de clase promedio para los estudiantes haciendo que todas las clases tengan un tamaño similar. Esto también aumentaría la asistencia, que es más alta en las clases con menor número de alumnos.

Lo que la mediana no es

El biólogo de Harvard, Stephen Jay Gould escribió: "La mediana no es el mensaje". Así describe lo que aprendió cuando se enteró que tenía mesotelioma abdominal, una forma de cáncer. Fue a la biblioteca para aprender más y se sorprendió al descubrir que el mesotelioma era incurable (la mediana de supervivencia es de sólo ocho meses después de ser descubierto). Gould escribió lo siguiente: "Sospecho que la mayor parte de las personas, sin formación en estadística, leería esta afirmación como 'probablemente estaré muerto en ocho meses'; conclusión que debe evitarse, ya que no es así, y más considerando que la actitud (en la lucha contra el cáncer) importa tanto". Gould procedió a interpretar cuidadosamente el valor de la mediana. Sabía que su oportunidad de vivir más tiempo que la mediana era buena porque era joven, su cáncer fue diagnosticado de manera temprana y tendría el mejor tratamiento médico. También razonó que algunas personas podrían vivir mucho más de ocho meses, y no vio ninguna razón por la que no pudiera estar en ese grupo. Armado con esta interpretación reflexiva de la mediana y una fuerte actitud positiva, Gould vivió 20 años después de su diagnóstico. Murió de otro cáncer no relacionado con el mesotelioma.

Mediana

La mediana se puede considerar de manera general como un "valor medio" en el sentido de que aproximadamente la mitad de los valores en un conjunto de datos son menores y la mitad son mayores que la mediana. La siguiente definición es más precisa.

DEFINICIÓN

La **mediana** de un conjunto de datos es la medida de tendencia central que indica el *valor intermedio*, cuando los datos originales se presentan en orden de magnitud creciente (o decreciente).

Propiedades importantes de la mediana

- La mediana no cambia por mucho cuando se incluyen sólo unos pocos valores extremos, por lo que la mediana es una medida de tendencia central *resistente*.
- La mediana no utiliza directamente todos los valores de datos. (Por ejemplo, si el valor mayor se cambia por uno mucho más grande, la mediana no cambia).

Cálculo y notación de la mediana

La mediana de una muestra se denomina a veces $\tilde{x}$ (que se pronuncia "x tilde") o M o Med; no existe una notación generalmente aceptada y no hay un símbolo especial para la mediana de una población. Para encontrar la mediana, primero *ordene* los valores (póngalos en orden) y luego siga uno de los siguientes dos procedimientos:

1. Si el número de valores de datos es *impar*, la mediana es el número ubicado en el intermedio exacto de la lista ordenada.
2. Si el número de valores de datos es *par*, la mediana se obtiene calculando la media de los dos números intermedios de la lista ordenada.

EJEMPLO 2 Mediana con un número *impar* de valores de datos

Encuentre la mediana de las cinco primeras velocidades de datos para Verizon: 38.5, 55.6, 22.4, 14.1 y 23.1 (todas en megabits por segundo, o Mbps).

SOLUCIÓN

Primero ordene en forma ascendente los valores de datos, como se muestra a continuación:

14.1 22.4 23.1 38.5 55.6

Como hay 5 valores de datos, tal cantidad es un número impar (5), por lo que la mediana es el número intermedio exacto de la lista ordenada, que es 23.1 Mbps. Por lo tanto, la mediana es 23.10 Mbps. Observe que la *mediana* de 23.10 Mbps es diferente de la *media* de 30.74 Mbps que se encontró en el ejemplo 1. Note también que el resultado de 23.10 Mbps sigue la regla de redondeo que se proporciona más adelante en esta sección.

SU TURNO Encuentre la mediana en el ejercicio 5 "Números de jugadores de fútbol americano".

EJEMPLO 3 Mediana con un número *par* de valores de datos

Repita el ejemplo 2 después de incluir la sexta velocidad de datos de 24.5 Mbps. Es decir, encuentre la mediana de las siguientes velocidades de datos: 38.5, 55.6, 22.4, 14.1, 23.1, 24.5 (todo en Mbps).

SOLUCIÓN

Primero ordene los valores de manera ascendente:

14.1 22.4 23.1 24.5 38.5 55.6

Debido a que el número de valores de datos es par (6), la mediana se obtiene calculando la media de los dos números intermedios, que son 23.1 y 24.5.

$$\text{Mediana} = \frac{23.1 + 24.5}{2} = \frac{47.6}{2} = 23.80 \text{ Mbps}$$

La mediana es 23.80 Mbps.

SU TURNO Encuentre la mediana en el ejercicio 7 "Valor neto de celebridades".

Moda

La moda no se utiliza mucho con datos cuantitativos, pero es la única medida de tendencia central que puede usarse con datos cualitativos (que consisten solamente en nombres, etiquetas o categorías).

En cifras

Mohammed: El nombre más común en el mundo.

DEFINICIÓN

La **moda** de un conjunto de datos es el (los) valor(es) que ocurre(n) con mayor frecuencia.

Propiedades importantes de la moda

- La moda se puede encontrar con datos cualitativos.
- Un conjunto de datos puede tener una moda, o múltiples modas, o no tener ninguna.

Determinación de la moda: Un conjunto de datos puede tener una moda, más de una moda, o ninguna moda.

- Cuando dos valores de datos ocurren con la misma mayor frecuencia, cada uno es una moda y se dice que el conjunto de datos es **bimodal**.
- Cuando más de dos valores de datos ocurren con la misma mayor frecuencia, cada uno es una moda y se dice que el conjunto de datos es **multimodal**.
- Cuando ningún valor de datos se repite, se dice que **no hay moda**.
- Cuando usted pide helado con su pastel, se dice que está "a la moda".

EJEMPLO 4 Moda

Encuentre la moda de las siguientes velocidades de datos para Sprint (en Mbps):

0.2 0.3 0.3 0.3 0.6 0.6 1.2

SOLUCIÓN

La moda es de 0.3 Mbps, porque es la velocidad de datos que ocurre con más frecuencia (tres veces).

SU TURNO Encuentre la moda en el ejercicio 7 "Valor neto de celebridades".

En el ejemplo 4, la moda es un valor único. A continuación se presentan otras circunstancias posibles:

Dos modas: Las velocidades de datos (Mbps) de 0.3, 0.3, 0.6, 4.0 y 4.0 tienen dos modas: 0.3 Mbps y 4.0 Mbps.

Sin moda: Las velocidades de datos (Mbps) de 0.3, 1.1, 2.4, 4.0 y 5.0 no tienen moda porque ningún valor se repite.

Mitad del rango

Otra medida de tendencia central es la mitad del rango.

DEFINICIÓN

La **mitad del rango** de un conjunto de datos es la medida de tendencia central que consiste en el valor que está a la mitad entre los valores máximo y mínimo del conjunto de datos original. Se encuentra al sumar el valor máximo y el valor mínimo de los datos y después dividir esa suma entre 2, como se muestra en la siguiente fórmula:

$$\text{Mitad del rango} = \frac{\text{valor máximo de datos} + \text{valor mínimo de datos}}{2}$$

Propiedades importantes de la mitad del rango

- Debido a que la mitad del rango utiliza sólo los valores máximo y mínimo, es muy sensible a esos extremos y, por lo tanto, *no es resistente*.
- En la práctica, la mitad del rango se utiliza con poca frecuencia, pero tiene tres características redentoras:
  1. Es muy fácil de calcular.
  2. Ayuda a reforzar la muy importante idea de que hay varias maneras de definir el centro de un conjunto de datos.
  3. En ocasiones, su valor se utiliza incorrectamente para la mediana, así que la confusión puede reducirse al definir claramente la mitad del rango junto con la mediana.

EJEMPLO 5 **Mitad del rango**

Encuentre la mitad del rango de las siguientes velocidades de datos para Verizon del ejemplo 1: 38.5, 55.6, 22.4, 14.1 y 23.1 (todas en Mbps)

SOLUCIÓN

La mitad del rango se encuentra de la siguiente manera:

$$\text{Mitad del rango} = \frac{\text{valor máximo de datos} + \text{valor mínimo de datos}}{2}$$

$$= \frac{55.6 + 14.1}{2} = 34.85 \text{ Mbps}$$

La mitad del rango es 34.85 Mbps.

SU TURNO Encuentre la mitad del rango en el ejercicio 5 “Números de jugadores de fútbol americano”.

Redondeo de medidas de tendencia central

A menudo, cuando se calculan medidas de tendencia central, es necesario redondear el resultado. Se utiliza la siguiente regla.

Reglas de redondeo para medidas de tendencia central

- **Para la media, la mediana y la mitad del rango, incluya un decimal más que los presentes en el conjunto original de valores.**
- **Para la moda, deje el valor sin redondear** (porque los valores de la moda son iguales que algunos de los valores de los datos originales).

Cuando aplique las reglas de redondeo, redondee sólo la respuesta final, *no los valores intermedios producidos durante los cálculos*. Por ejemplo, la media de 2, 3 y 5 es 3.333333..., que se redondea a 3.3; esto es, con un decimal más que los valores originales de 2, 3 y 5. Otro ejemplo es: la media de 80.4 y 80.6 es 80.50 (un decimal más que los usados para los valores originales). Debido a que la moda es uno o más de los valores originales, no la redondeamos; simplemente usamos los mismos valores originales que resultaron ser modas.

Pensamiento crítico

Siempre es posible calcular las medidas de tendencia central a partir de una muestra de números, pero es necesario considerar si hacerlo tiene sentido. En la sección 1-2 se estableció que no tiene sentido hacer cálculos numéricos con datos al nivel nominal de medición, ya que estos datos constan sólo de nombres, etiquetas o categorías, por lo que los datos estadísticos como la media y la mediana carecen de significado. También se debe pensar en el método de muestreo utilizado para recopilar los datos. Si el método de muestreo no es sólido, los datos estadísticos obtenidos pueden ser muy engañosos.

EJEMPLO 6 **Pensamiento crítico y medidas de tendencia central**

Considere cada una de las siguientes situaciones ilustrativas en las que la media y la mediana *no* son estadísticos significativos.

a. Códigos postales del Gateway Arch en San Luis, la Casa Blanca, la División de la Fuerza Aérea del Pentágono, el Edificio Empire State y la Estatua de la Libertad: 63102, 20500, 20330, 10118, 10004. (Los códigos postales no miden ni cuentan nada. Los números son sólo etiquetas para ubicaciones geográficas).

b. Clasificaciones de universidades nacionales seleccionadas: Harvard, Yale, Duke, Dartmouth y Brown (por *US News & World Report*): 2, 3, 7, 10, 14. (Las clasificaciones reflejan una ordenación, pero no miden ni cuentan nada).

c. Números en las camisetas de la defensiva titular de los Halcones marinos de Seattle cuando ganaron el Súper Tazón XLVIII: 31, 28, 41, 56, 25, 54, 69, 50, 91, 72, 29. (Los números en las camisetas de fútbol americano no miden ni cuentan nada, son sólo sustitutos de los nombres).

d. Primeros cinco ingresos de directores de empresa (en millones de dólares): 131.2, 66.7, 64.4, 53.3, 51.5. (Estas listas de "los primeros 5" o "los primeros 10" incluyen datos que no son en absoluto representativos de toda la población).

e. Las 50 edades medias calculadas a partir de las medias en cada uno de los 50 estados. (Si se calcula la media de esos 50 valores, el resultado no es la edad media de las personas en todo el territorio de Estados Unidos y se debe tener en cuenta los tamaños de la población de los 50 estados, como se describe en la introducción de la *media ponderada* en la parte 2 de esta sección).

SU TURNO Para el ejercicio 5 "Números de jugadores de fútbol americano", determine por qué la media y la mediana no son significativas.

Con la idea de describir, explorar y comparar datos, proporcionamos la tabla 3-1 en la página 88, que resume las diferentes medidas de tendencia central para las velocidades de datos para teléfonos inteligentes que se mencionan en el problema del capítulo. Los datos se listan en el conjunto de datos 32 "Velocidades de datos en aeropuertos" del apéndice B. La figura 3-1 de la página 81 sugiere que Verizon tiene las mayores velocidades y la comparación de medias y medianas en la tabla 3-1 también lo indica. En los siguientes capítulos se describen otras herramientas que pueden utilizarse para una comparación efectiva.

Un error de redondeo cambia un récord mundial

Los errores de redondeo a menudo pueden tener resultados desastrosos. Justin Gatlin estaba eufórico cuando estableció el récord mundial como la persona en correr los 100 metros en el menor tiempo (9.76 segundos). Sin embargo, su tiempo récord duró sólo cinco días, cuando se corrigió a 9.77 segundos y empató el récord mundial en lugar de romperlo. Su tiempo real fue de 9.766 segundos, y debería haberse redondeado a 9.77 segundos, pero la persona que tomó el tiempo no sabía que tenía que presionar un botón para obtener el redondeo. El representante de Gatlin dijo que el atleta estaba muy perturbado y que el incidente era una "gran vergüenza para la IAAF (International Association of Athletics Federations) y para nuestro deporte".

TABLA 3-1 Comparación de las velocidades de datos para teléfonos inteligentes (Mbps) en los aeropuertos

	Verizon	Sprint	AT&T	T-Mobile
Media	17.60	3.71	10.70	10.99
Mediana	13.90	1.60	8.65	9.70
Moda	4.5, 11.1	0.3	2.7	3.2, 4.4, 5.1, 13.3, 15.0, 16.7, 27.3
Mitad del rango	39.30	15.30	19.80	14.00

PARTE 2 Más allá de lo básico en las medidas de tendencia central

Cálculo de la media a partir de una distribución de frecuencias

La fórmula 3-2 es el mismo cálculo para la media que se presentó en la parte 1, pero incorpora el siguiente método: cuando trabajamos con datos resumidos en una distribución de frecuencias, hacemos posibles los cálculos asumiendo que todos los valores muestrales en cada clase son iguales al punto medio de dicha clase. La fórmula 3-2 no es realmente un concepto nuevo; es simplemente una variación de la fórmula 3-1 (media).

FÓRMULA 3-2 MEDIA A PARTIR DE UNA DISTRIBUCION DE FRECUENCIAS

Primero multiplique cada frecuencia y el punto medio de la clase; luego sume los productos.

$$\bar{x} = \frac{\Sigma(f \cdot x)}{\Sigma f}$$

(El resultado es una aproximación)

Suma de frecuencias (igual a n)

El ejemplo 7 ilustra el procedimiento para encontrar la media a partir de una distribución de frecuencias.

EJEMPLO 7 Cálculo de la media a partir de una distribución de frecuencias

Las dos primeras columnas de la tabla 3-2 mostradas aquí son las mismas de la distribución de frecuencias de la tabla 2-2 del capítulo 2. Utilice la distribución de frecuencias en las dos primeras columnas de la tabla 3-2 para encontrar la media.

TABLA 3-2 Tiempos de servicio para el almuerzo en McDonald's

Tiempo (segundos)	Frecuencia f	Punto medio de la clase x	$f \cdot x$
75-124	11	99.5	1094.5
125-174	24	149.5	3588.0
175-224	10	199.5	1995.0
225-274	3	249.5	748.5
275-324	2	299.5	599.0
Totales:	$\Sigma f = 50$		$\Sigma(f \cdot x) = 8025.0$

SOLUCIÓN

Recuerde que al trabajar con datos resumidos en una distribución de frecuencias, se hacen posibles los cálculos asumiendo que todos los valores muestrales en cada clase son iguales al punto medio de dicha clase. Por ejemplo, considere el intervalo de la primera clase de 75 a 124 con una frecuencia de 11. Asumimos que cada uno de los 11 tiempos de servicio es de 99.5 segundos (el punto medio de la clase). Con el tiempo de servicio de 99.5 segundos repetido 11 veces, tenemos un total de $99.5 \cdot 11 = 1094.5$, como se muestra en la última columna de la tabla 3-2. Después sumamos los resultados para encontrar la sumatoria de todos los valores muestrales.

La fila inferior de la tabla 3-2 muestra los dos componentes que se requieren para calcular la media (como en la fórmula 3-2) $\Sigma f = 50$ y $\Sigma (f \cdot x) = 8025.0$. Calculamos la media usando la fórmula 3-2 como sigue:

$$\bar{x} = \frac{\Sigma (f \cdot x)}{\Sigma f} = \frac{8025.0}{50} = 160.5 \text{ segundos}$$

El resultado de $\bar{x} = 160.5$ segundos es una *aproximación* porque se basa en el uso de los valores de los puntos medios de las clases en lugar de la lista original de tiempos de servicio. La media de 160.2 segundos encontrada mediante el uso de todos los tiempos de servicio originales es un resultado más preciso.

SU TURNO Resuelva el ejercicio 29 “Edades de las mejores actrices”.

Cálculo de una media ponderada

Cuando a los diferentes valores de datos x se les asignan pesos w distintos, podemos calcular la media ponderada, que está dada por la fórmula 3-3.

FÓRMULA 3-3

$$\text{Media ponderada: } \bar{x} = \frac{\Sigma (w \cdot x)}{\Sigma w}$$

La fórmula 3-3 implica primero multiplicar cada peso w por el valor correspondiente x, luego sumar los productos, y finalmente dividir el total por la suma de los pesos, Σw.

EJEMPLO 8 Cálculo del promedio de calificaciones

En su primer semestre de la universidad, una alumna del autor tomó cinco cursos. Sus calificaciones finales, junto con el número de créditos para cada curso, fueron A (3 créditos), A (4 créditos), B (3 créditos), C (3 créditos) y F (1 crédito). El sistema de clasificación asigna puntos de calidad a las calificaciones con letras de la siguiente manera: A = 4; B = 3; C = 2; D = 1; F = 0. Calcule su promedio de calificaciones.

SOLUCIÓN

Use los números de créditos como pesos: $w = 3, 4, 3, 3, 1$. Reemplace las calificaciones con letras de A, A, B, C y F con los puntos de calidad correspondientes: $x = 4, 4, 3, 2, 0$. Ahora usamos la fórmula 3-3 como se muestra a continuación. El resultado es un promedio de 3.07 en el primer semestre. (Al utilizar la regla de redondeo precedente, el resultado debe redondearse a 3.1, pero es común que los promedios de calificaciones se redondeen a dos decimales).

$$\begin{aligned} \bar{x} &= \frac{\Sigma (w \cdot x)}{\Sigma w} \\ &= \frac{(3 \times 4) + (4 \times 4) + (3 \times 3) + (3 \times 2) + (1 \times 0)}{3 + 4 + 3 + 3 + 1} \\ &= \frac{43}{14} = 3.07 \end{aligned}$$

SU TURNO Resuelva el ejercicio 33 “Media ponderada”.

CENTRO DE TECNOLOGÍA

Ejemplos de pantallas de estadística descriptiva

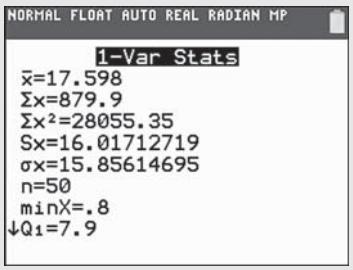
Acceda a los complementos de software, videos y conjuntos de datos en www.pearsonenespañol.com/triola

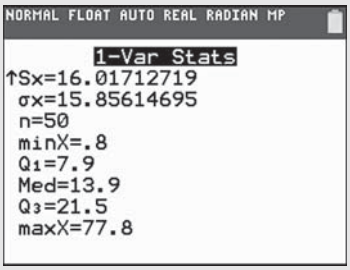
Las siguientes pantallas se basan en las velocidades de datos de Verizon dentro del conjunto de datos 32 “Velocidades de datos en aeropuertos”.

Statdisk

```
Explore Data - Column 2

Sample Size, n: 50
Mean: 17.598
Median: 13.9
Midrange: 39.3
RMS: 23.6877
Variance, s^2: 256.5484
Standard Deviation, s: 16.01713
Mean Absolute Deviation: 10.66528
Range: 77
Coefficient of Variance: 91.02%

Minimum: 0.8
1st Quartile: 7.9
2nd Quartile: 13.9
3rd Quartile: 21.5
Maximum: 77.8

Sum: 879.9
Sum of Squares: 28055.35

95% CI for the Mean:
13.046 < mean <22.15

95% CI for the Standard Deviation:
13.3796 < SD < 19.9595

95% CI for the Variance:
179.0149 < VAR < 398.3804
```

TI-83/4 Plus





Minitab

Descriptive Statistics: VERIZON

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
VERIZON	50	0	17.60	2.27	16.02	0.80	7.85	13.90	21.68	77.80

StatCrunch

Summary statistics:

Column	n	Mean	Variance	Std. dev.	Std. err.	Median	Range	Min	Max	Q1	Q3
VERIZON	50	17.598	256.54836	16.017127	2.2651638	13.9	77	0.8	77.8	7.9	21.5

Herramienta de análisis de datos de Excel

VERIZON	
Mean	17.598
Standard Error	2.26516385
Median	13.9
Mode	11.1
Standard Deviation	16.01712719
Sample Variance	256.5483633
Kurtosis	5.485898877
Skewness	2.193633453
Range	77
Minimum	0.8
Maximum	77.8
Sum	879.9
Count	50

Suplemento XLSTAT Excel

Statistic	VERIZON
Nbr. of observations	50
Minimum	0.800
Maximum	77.800
Freq. of minimum	1
Freq. of maximum	1
Range	77.000
1st Quartile	8.400
Median	13.900
3rd Quartile	21.400
Sum	879.900
Mean	17.598
Variance (n)	251.417
Variance (n-1)	256.548
Standard deviation (n)	15.856
Standard deviation (n-1)	16.017

CENTRO DE TECNOLOGÍA

Estadística descriptiva

Acceda a los complementos tecnológicos, videos y conjuntos de datos de Elementary Statistic, 13E en www.pearsonenespañol.com/triola

Statdisk

1. Haga clic en **Data** en el menú superior.
2. Seleccione **Explore Data-Descriptive Statistics** en el menú desplegable.
3. Seleccione la columna de datos deseada.
4. Haga clic en **Evaluate** para ver los estadísticos descriptivos.

Minitab

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Basic Statistics** en el menú desplegable y después elija **Display Descriptive Statistics**.
3. Haga doble clic en la columna de datos deseada para que aparezca en la ventana *Variables*.
4. Haga clic en **OK** para ver los estadísticos descriptivos.

SUGERENCIA: Haga clic en el botón **Statistics** arriba de **OK** para seleccionar los estadísticos individuales que desea mostrar.

StatCrunch

1. Haga clic en **Stat** en el menú superior.
2. Seleccione **Summary Stats** en el menú desplegable y después elija **Columns**.
3. Seleccione la columna de datos deseada.
4. Haga clic en **Compute!** para ver los estadísticos descriptivos.

SUGERENCIA: Personalice los estadísticos descriptivos al seleccionar los elementos bajo **Statistics**.

Calculadora T1-83/84 Plus

1. Presione STAT, luego seleccione **CALC** en el menú superior.
2. Seleccione **1-Var Stats** y pulse ENTER.
3. Introduzca el nombre de la lista que incluye los datos deseados (por ejemplo, L1).
4. Seleccione **Calculate** y pulse ENTER para ver los estadísticos descriptivos.

SUGERENCIA: Presione ▽ para ver los estadísticos adicionales que no se desplegaron en la pantalla inicial.

Excel

Complemento XLSTAT

1. Haga clic en la ficha **XLSTAT** de la barra de opciones y después haga clic en **Describing Data**.
2. Seleccione **Descriptive Statistics** en el menú desplegable.
3. Marque la casilla **Quantitative Data** e introduzca el intervalo de datos deseado. Si la primera fila de datos contiene una etiqueta, también marque la casilla **Sample labels**.
4. Haga clic en **OK** para ver los estadísticos descriptivos.

Complemento Excel Data Analysis

1. Haga clic en la ficha **Data** de la barra de opciones y después seleccione **Data Analysis** en el menú superior.
2. Seleccione **Descriptive Statistics** bajo *Analysis Tools*.
3. Introduzca el rango de datos deseado para el **Inpute Range**. Si la primera fila de datos contiene una etiqueta, también marque la casilla **Labels in First Row**.
4. Marque la etiqueta **Summary Statistics** y haga clic en **OK** para ver los estadísticos descriptivos.

3-1 Habilidades y conceptos básicos

Conocimiento estadístico y pensamiento crítico

1. Promedio El extinto sitio web IncomeTaxlist.com indicaba que el ingreso anual "promedio" en Florida era de $35,031. ¿Cuál es el papel del término *promedio* en estadística? ¿Debería usarse otro término en lugar de *promedio*?

2. ¿Qué hay de erróneo? *USA Today* publicó una lista consistente en el impuesto estatal por cada galón de gasolina. Si se suman las 50 cantidades de impuestos estatales y luego se divide la suma por 50, se obtiene 27.3 centavos. ¿Es el valor de 27.3 centavos el importe medio del impuesto estatal sobre ventas pagado por todos los conductores estadounidenses? ¿Por qué sí o por qué no?

3. Medidas de tendencia central ¿En qué sentido la media, la mediana, la moda y la mitad del rango son medidas del "centro"?

4. Medidas resistentes Las siguientes son cuatro de las velocidades de datos Verizon (Mbps) de la figura 3-1: 13.5, 10.2, 21.1, 15.1. Encuentre la media y la mediana de estos cuatro valores. A continuación, determine la media y la mediana después de incluir un quinto valor de 142, que es un valor atípico. (Una de las velocidades de datos de Verizon es 14.2 Mbps, pero 142 se utiliza aquí como un error resultante de una entrada con un punto decimal faltante). Compare los dos conjuntos de resultados. ¿Qué

tanto se vio afectada la media por la inclusión del valor atípico? ¿Cuánto fue afectada la mediana por la inclusión del valor atípico?

Pensamiento crítico. ***En los ejercicios 5 a 20, tenga cuidado con las pequeñas trampas. Cada uno de estos ejercicios implica alguna característica capciosa. Encuentre la (a) media (b) mediana, (c) moda, (d) mitad del rango, y luego responda la pregunta dada.***

5. Números de jugadores de fútbol americano A continuación se listan los números de camiseta de 11 jugadores seleccionados al azar del equipo de los Halcones marinos de Seattle, cuando ganaron el Súper Tazón XLVIII. ¿Qué nos dicen los resultados?

89 91 55 7 20 99 25 81 19 82 60

6. Pesos de jugadores de fútbol americano A continuación se listan los pesos en libras de 11 jugadores seleccionados al azar del equipo de los Halcones marinos de Seattle, cuando ganaron el Súper Tazón XLVIII (los mismos jugadores del ejercicio anterior). ¿Es probable que los resultados sean representativos de todos los jugadores de la NFL?

189 254 235 225 190 305 195 202 190 252 305

7. Valor neto de celebridades A continuación se listan los mayores valores netos (en millones de dólares) de las celebridades. Las celebridades son Tom Cruise, Will Smith, Robert De Niro, Drew Carey, George Clooney, John Travolta, Samuel L. Jackson, Larry King, Demi Moore y Bruce Willis. ¿Qué nos dicen los resultados sobre la población de todas las celebridades? Sobre la base de la naturaleza de las cantidades, ¿qué se puede inferir acerca de su precisión?

250 200 185 165 160 160 150 150 150 150

8. Lo que pasa en Las Vegas... A continuación se listan los precios en dólares por una noche en diferentes hoteles ubicados en Las Vegas Boulevard (the "Strip"). Si usted decide hospedarse en uno de estos hoteles, ¿qué estadístico es más relevante, además de las medidas de tendencia central? Aparte del precio, identifique otro factor importante que afectaría su elección.

212 77 121 104 153 264 195 244

9. Huracanes A continuación se listan las cantidades de huracanes que se produjeron en el Atlántico cada año. Los datos se dan en orden anual, a partir del año 2000. ¿Qué característica importante de los datos no es revelada por ninguna de las medidas de tendencia central?

8 9 8 7 9 15 5 6 8 4 12 7 8 2

10. Chícharos en una vaina Los biólogos han realizado experimentos para determinar si una falta de bióxido de carbono en el suelo afecta a los fenotipos de los chícharos. A continuación se listan los códigos del fenotipo, donde 1 = amarillo liso, 2 = verde liso, 3 = amarillo corrugado y 4 = verde corrugado. ¿Se pueden obtener las medidas de tendencia central para estos valores? ¿Los resultados tienen sentido?

2 1 1 1 1 1 1 4 1 2 2 1 2 3 3 2 3 1 3 1 3 1 3 2 2

11. Precios de televisores A continuación se listan los precios de venta (en dólares) de televisores de 60 pulgadas o más, los cuales fueron calificados como "las mejores compras" por la revista *Consumer Reports*. ¿Los datos estadísticos resultantes son representativos de la población de todos los televisores de 60 pulgadas o más? Si usted decide comprar uno de estos televisores, ¿qué dato estadístico es más relevante, además de las medidas de tendencia central?

1800 1500 1200 1500 1400 1600 1500 950 1600 1150 1500 1750

12. Radiación del teléfono celular A continuación se listan las tasas de absorción de radiación medidas (en W/kg) correspondientes a los siguientes teléfonos celulares: iPhone 5S, BlackBerry Z30, Sanyo Vero, Optimus V, Droid Razr, Nokia N97, Samsung Vibrant, Sony Z750a, Kyocera Kona, LG G2 y Virgin Mobile Supreme. Los datos provienen de la Comisión Federal de Comunicaciones (FCC, por sus siglas en inglés). Los medios de comunicación a menudo informan sobre los peligros de la radiación de los teléfonos celulares como una causa de cáncer. La FCC tiene un estándar de que la tasa de absorción de un teléfono celular debe ser de 1.6 W/kg o menos. Si usted está planeando comprar un teléfono celular, ¿es alguna de las medidas de tendencia central el estadístico más importante? ¿Hay otro dato estadístico que sea más relevante? Si es así, ¿cuál?

1.18 1.41 1.49 1.04 1.45 0.74 0.89 1.42 1.45 0.51 1.38

13. Cafeína en bebidas A continuación se indican las cantidades medidas de cafeína (mg por 12 onzas de bebida) obtenidas en una lata de cada una de 20 marcas (7-UP, A & W Root Beer, Cherry Coke, ...). ¿Son los estadísticos representativos de la población de todas las latas de las mismas 20 marcas consumidas por los estadounidenses?

0 0 34 34 34 45 41 51 55 36 47 41 0 0 53 54 38 0 41 47

14. Muertes de bomberos A continuación se listan las cantidades de heroicos bomberos que perdieron sus vidas en Estados Unidos anualmente mientras combatían incendios forestales. Las cantidades están ordenadas por año, comenzando en 2000. ¿Qué característica importante de los datos no es revelada por ninguna de las medidas de tendencia central?

20 18 23 30 20 12 24 9 25 15 8 11 15 34

15. Longitud de los pies A continuación se listan las longitudes en pulgadas de los pies de mujeres del ejército seleccionadas al azar, medidos en el sondeo antropométrico de 1988 (ANSUR, abreviatura en inglés). Estos datos estadísticos, ¿son representativos de la población actual de todas las mujeres del ejército?

10.4 9.3 9.1 9.3 10.0 9.4 8.6 9.8 9.9 9.1 9.1

16. Universidades más caras A continuación se listan los costos anuales (en dólares) de matrícula y colegiaturas en las 10 universidades más caras de Estados Unidos para un año reciente (con base en datos de *US News & World Report*). Las universidades listadas en orden son Columbia, Vassar, Trinidad, George Washington, Carnegie Mellon, Wesleyan, Tulane, Bucknell, Oberlin y Union. ¿Qué nos dice esta "lista de las 10 primeras" sobre esos costos para la población de todas las universidades estadounidenses?

49,138 47,890 47,510 47,343 46,962 46,944 46,930 46,902 46,870 46,785

17. Anillo de diamante A continuación se listan las cantidades en dólares que cuestan los paquetes para propuestas matrimoniales en los diferentes estadios de las Ligas Mayores de Béisbol. Cinco de los equipos no permiten propuestas. ¿Existen valores atípicos?

39 50 50 50 55 55 75 85 100 115 175 175 200
209 250 250 350 400 450 500 500 500 500 1500 2500

18. Ventas de álbumes de discos LP de vinilo A continuación se listan las ventas anuales de discos de vinilo en Estados Unidos (millones de unidades). Las cantidades de álbumes vendidos se presentan en orden cronológico, y la última entrada representa el año más reciente. ¿Las medidas de tendencia central nos dan alguna información sobre una tendencia cambiante a lo largo del tiempo?

0.3 0.6 0.8 1.1 1.1 1.4 1.4 1.5 1.2 1.3 1.4 1.2 0.9 0.9
1 1.9 2.5 2.8 3.9 4.6 6.1

19. Fumadores de California En la encuesta Entrevista sobre Salud en California, se entrevistó a adultos seleccionados al azar. Una de las preguntas fue acerca de cuántos cigarrillos fumaban al día, y los resultados se listan a continuación para 50 encuestados seleccionados aleatoriamente. ¿Qué tan bien reflejan los resultados los hábitos del tabaquismo de los adultos en California?

9 10 10 20 40 50 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0

20. Citas rápidas En un estudio sobre citas rápidas realizado en la Universidad de Columbia, se pidió a las mujeres que evaluaran el atractivo de sus acompañantes masculinos; a continuación se lista una muestra de los resultados (1 = no atractivo, 10 = extremadamente atractivo). ¿Se pueden utilizar los resultados para describir el atractivo de la población de varones adultos?

5 8 3 8 6 10 3 7 9 8 5 5 6 8 8 7 3 5 5 6 8 7 8 8 8 7

En los ejercicios 21 a 24, encuentre la media y la mediana para cada una de las dos muestras, luego compare los dos conjuntos de resultados.

21. Presión arterial Una muestra de mediciones de presión arterial se toma del conjunto de datos 1 "Datos corporales" del apéndice B, y los valores (mm Hg) se relacionan de manera que 10 sujetos tienen medidas sistólicas y diastólicas. (La presión sistólica es una medida de la fuerza de la sangre empujada a través de las arterias, mientras que la presión diastólica es una medida de la presión arterial mientras el corazón está en reposo entre latidos). ¿Son las medidas de tendencia central los mejores estadísticos que pueden obtenerse de estos datos?

Sistólica:	118	128	158	96	156	122	116	136	126	120
Diastólica:	80	76	74	52	90	88	58	64	72	82

22. Robo de parquímetros A continuación se listan los montos (en millones de dólares) recaudados en parquímetros por Brinks y otras empresas en la ciudad de Nueva York durante períodos similares. Se utilizó un conjunto de datos más amplio para condenar a cinco empleados de Brinks por hurto mayor. Los datos fueron proporcionados por el abogado de la ciudad de Nueva York, y se pueden encontrar en el sitio web de *Data and Story Library* (DASL). ¿Los datos limitados que aparecen aquí muestran evidencia de robo por los empleados de Brinks?

El contratista de recaudación fue Brinks	1.3	1.5	1.3	1.5	1.4	1.7	1.8	1.7	1.7	1.6
El contratista de recaudación no fue Brinks	2.2	1.9	1.5	1.6	1.5	1.7	1.9	1.6	1.6	1.8

23. Pulsos A continuación se listan los pulsos (latidos por minuto) de muestras de hombres y mujeres adultas (del conjunto de datos 1 "Datos corporales" en el apéndice B). ¿Parece haber una diferencia?

Hombres:	86	72	64	72	72	54	66	56	80	72	64	64	96	58	66
Mujeres:	64	84	82	70	74	86	90	88	90	90	94	68	90	82	80

24. Filas en el banco Los tiempos de espera (en segundos) de los clientes en el Banco de Ahorro de Madison se registran con dos configuraciones: línea de clientes única y líneas de clientes individuales. Examine cuidadosamente los datos para determinar si hay una diferencia entre los dos conjuntos de datos que no sea evidente a partir de una comparación de las medidas de tendencia central. Si es así, ¿cuál es?

Línea única	390	396	402	408	426	438	444	462	462	462
Líneas individuales	252	324	348	372	402	462	462	510	558	600

Grandes conjuntos de datos del apéndice B. ***En los ejercicios 25 a 28, considere el conjunto de datos indicado en el apéndice B. Use software o una calculadora para encontrar las medias y las medianas.***

25. Tornados Utilice las medidas en la escala F de tornados listadas en el conjunto de datos 22 "Tornados" del apéndice B. Entre los 500 tornados, ¿a cuántos les faltan mediciones de la escala F? (*Precaución:* En algunas tecnologías, los datos faltantes se representan mediante una constante como –9 o 9999).

26. Terremotos Utilice las magnitudes (en la escala de Richter) de los 600 terremotos listados en el conjunto de datos 21 "Terremotos" del apéndice B. En 1989, el área de la Bahía de San Francisco fue golpeada por un terremoto de 7.0 en la escala de Richter. Ese terremoto ocurrió durante el período de calentamiento para el tercer juego de la Serie Mundial de béisbol. ¿Es la magnitud del terremoto de la Serie Mundial un valor atípico cuando se considera en el contexto de los datos muestrales dados en el conjunto de datos 21? Explique.

27. Temperaturas corporales Considere el conjunto de datos 3 "Temperaturas corporales" en el apéndice B y use las temperaturas corporales para las 12:00 AM del día 2. ¿Los resultados respaldan o contradicen la creencia común de que la temperatura corporal media es de 98.6 °F?

28. Nacimientos Use los pesos al nacer (en gramos) de los 400 bebés listados en el conjunto de datos 4 "Nacimientos" en el apéndice B. Examine la lista de pesos al nacer y haga una observación sobre esos números. ¿Cómo afecta esta observación la forma en que los resultados deben redondearse?

En los ejercicios 29 a 32, encuentre la media de los datos resumidos en la distribución de frecuencias. Además compare las medias calculadas con las medias reales obtenidas utilizando la lista original de valores de datos, que son las siguientes: (Ejercicio 29) 36.2 años; (Ejercicio 30) 44.1 años; (Ejercicio 31) 224.3; (Ejercicio 32) 255.1.

29.

Edad (en años) de la mejor actriz al ganar el Oscar	Frecuencia
20-29	29
30-39	34
40-49	14
50-59	3
60-69	5
70-79	1
80-89	1

30.

Edad (en años) del mejor actor al ganar el Oscar	Frecuencia
20-29	1
30-39	28
40-49	36
50-59	15
60-69	6
70-79	1

31.

Conteo de plaquetas en la sangre de hombres (1000 células/μl)	Frecuencia
0-99	1
100-199	51
200-299	90
300-399	10
400-499	0
500-599	0
600-699	1

32.

Conteo de plaquetas en la sangre de mujeres (1000 células/μl)	Frecuencia
100-199	25
200-299	92
300-399	28
400-499	0
500-599	2

33. Media ponderada Un alumno del autor obtuvo calificaciones de A, C, B, A y D. Estos cursos tenían las siguientes cantidades correspondientes de horas de crédito: 3, 3, 3, 4 y 1. El sistema de calificación asigna puntos a las calificaciones con letras como sigue: A = 4; B = 3; C = 2; D = 1; F = 0. Calcule el promedio de calificaciones y redondee el resultado con dos decimales. Si la lista del decano requiere un promedio de 3.00 o mayor, ¿este estudiante entrará a la lista del decano?

34. Media ponderada Una alumna del autor obtuvo calificaciones de 63, 91, 88, 84 y 79 en sus cinco exámenes regulares de estadística. Su calificación en el examen final fue 86 y en su proyecto de clase obtuvo 90. Su calificación combinada en las tareas fue 70. Los cinco exámenes regulares representan 60% de la calificación final, el examen final 10%, el proyecto 15% y las tareas 15%. ¿Cuál es su calificación media ponderada? ¿Qué calificación con letra obtuvo (A, B, C, D o F)? Suponga que una media de 90 o más es una A, una media de 80 a 89 es una B, y así sucesivamente.

3-1 Más allá de lo básico

35. Grados de Libertad Cinco frecuencias de pulso aleatoriamente seleccionadas del conjunto de datos 1 "Datos corporales" en el apéndice B tienen una media de 78.0 latidos por minuto. Cuatro de los pulsos son 82, 78, 56 y 84.

a. Encuentra el valor faltante.

b. Necesitamos elaborar una lista de n valores que tengan una media específica conocida. Somos libres de seleccionar los valores que deseemos para algunos de los n valores. ¿Cuántos de los n valores pueden asignarse libremente antes de que los valores restantes queden determinados? (El resultado se conoce como el *número de grados de libertad*).

36. Datos censurados El conjunto de datos 15 "Presidentes" del apéndice B muestra el número de años que vivieron los presidentes estadounidenses después de su primera toma de poder. Hasta la fecha, cinco de los presidentes siguen vivos y después de su primera asunción han vivido 37, 25, 21, 13 y 5 años. Podemos usar los valores de 37+, 25+, 21+, 13+ y 5+, donde los signos positivos indican

que el valor real es igual o mayor que el valor actual. (Se dice que estos valores están *censurados* en el momento actual en que se compiló a lista). Si se usan los valores en el conjunto de datos 15 y se ignoran los presidentes que todavía están vivos, ¿cuál es la media? Si utiliza los valores dados en el conjunto de datos 15 junto con los valores adicionales de 37+, 25+, 21+, 13+ y 5+, ¿qué sabemos acerca de la media? ¿Los dos resultados difieren mucho?

37. Media recortada Debido a que la media es muy sensible a los valores extremos, se dice que no es una medida de tendencia central resistente. Al eliminar algunos valores bajos y altos, la **media recortada** se hace más resistente. Para encontrar la media recortada del 10% para un conjunto de datos, primero ordene los datos, luego elimine el 10% de los valores inferiores y el 10% de los valores superiores, luego calcule la media de los valores restantes. Utilice las cargas axiales (en libras) de las latas de aluminio que se listan a continuación (del conjunto de datos 30 "Latas de aluminio" en el apéndice B) para latas de 0.0111 pulgadas de espesor. Una carga axial es la fuerza a la que la tapa de una lata colapsa. Identifique cualquier valor atípico, luego compare la mediana, la media, la media recortada del 10% y la media recortada del 20%.

247 260 268 273 276 279 281 283 284 285 286 288
289 291 293 295 296 299 310 504

38. Media armónica La **media armónica** se utiliza a menudo como una medida de tendencia central para conjuntos de datos que consisten en tasas de cambio, como velocidades. Se encuentra al dividir el número de valores n por la suma de los recíprocos de todos los valores, expresados como

$$\frac{n}{\sum \frac{1}{x}}$$

(Ningún valor puede ser cero). El autor condujo 1163 millas para ir a una conferencia en Orlando, Florida. En el viaje de ida, el autor se detuvo durante la noche, y la velocidad media de principio a fin fue de 38 millas por hora. Durante el viaje de regreso, se detuvo sólo por comida y combustible, y la velocidad media desde el principio hasta el final fue de 56 millas por hora. Encuentre la media armónica de 38 mi/h y 56 mi/h para encontrar la verdadera velocidad "media" del viaje de ida y vuelta.

39. Media geométrica La **media geométrica** se utiliza a menudo en los negocios y la economía para encontrar tasas de cambio promedio, tasas de crecimiento promedio o razones medias. Para encontrar la media geométrica de n valores (donde todos son positivos), primero multiplique los valores, luego encuentre la raíz n-ésima del producto. Por un período de 6 años, el dinero depositado en certificados anuales de depósito tenía tasas de interés anual de 5.154%, 2.730%, 0.488%, 0.319%, 0.313% y 0.268%. Identifique el porcentaje de crecimiento único que es igual que las cinco tasas de crecimiento consecutivas, calculando la media geométrica de 1.05154, 1.02730, 1.00488, 1.00319, 1.00313 y 1.00268.

40. Media cuadrática La **media cuadrática** (o la **raíz cuadrada media**, o **R.C.M.**) se utiliza en aplicaciones físicas, como sistemas de distribución de energía. La media cuadrática de un conjunto de valores se obtiene al elevar al cuadrado cada valor, sumar esos cuadrados, dividir la suma por el número de valores n, y luego obtener la raíz cuadrada del resultado, como se indica a continuación:

$$\text{Media cuadrática} = \sqrt{\frac{\Sigma x^2}{n}}$$

Encuentre la R.C.M. de los siguientes voltajes medidos en corrientes domésticas: 0, 60, 110, –110, –60, 0. ¿Cómo se compara el resultado con la media?

41. Mediana Cuando los datos se resumen en una distribución de frecuencias, la mediana se puede encontrar identificando primero la clase mediana, que es la clase que contiene la mediana. Asumimos entonces que los valores de esa clase están uniformemente distribuidos e interpolamos. Si n expresa la suma de todas las frecuencias de clase, y m expresa la suma de las frecuencias de clase que preceden a la clase mediana, la mediana se puede estimar como se muestra a continuación.

$$(\text{límite inferior de la clase mediana}) + (\text{anchura de clase})\left(\frac{\left(\frac{n+1}{2}\right) - (m+1)}{\text{frecuencia de la clase mediana}}\right)$$

Utilice este procedimiento para encontrar la mediana de la distribución de frecuencias dada en la tabla 3-2 de la página 88. ¿Cuánto se aleja este resultado de la mediana encontrada en la lista original de tiempos de servicio en McDonald's del conjunto de datos 25 "Comida rápida" en el apéndice B?

3-2 Medidas de variación

Concepto clave La variación es el tema más importante en estadística, por lo que ésta es la sección más importante del libro. La presente sección introduce tres medidas importantes de variación: el *rango*, la *desviación estándar* y la *varianza*. Estos estadísticos son números, pero nuestro enfoque no es sólo calcular esos números, sino desarrollar la capacidad de *interpretarlos* y *comprenderlos*. Esta sección no es un estudio de aritmética; se trata de entender e interpretar medidas de variación, en especial la desviación estándar.

SUGERENCIA DE ESTUDIO La parte 1 de esta sección introduce conceptos básicos de la variación, y la parte 2 presenta conceptos adicionales relacionados con la desviación estándar. Ambas partes incluyen fórmulas para el cálculo, pero no dedican mucho tiempo a la enseñanza de fórmulas o la realización de cálculos aritméticos. En cambio, se centran en *entender* e *interpretar* los valores de la desviación estándar.

PARTE 1 Conceptos básicos de variación

Para visualizar la propiedad de variación, vea la figura 3-2, que ilustra los tiempos de espera (segundos) de los clientes de un banco en dos condiciones diferentes: (1) Todos los clientes ingresan a una sola fila que alimenta a diferentes cajeros; (2) todos los clientes eligen unirse a la línea en uno de *varios* cajeros. Verifique esta observación importante: los tiempos de espera con la línea única (gráfica superior) tienen menos *variación* que los tiempos de espera con múltiples líneas (gráfica inferior). *Ambos conjuntos de tiempos de espera tienen la misma media de 100.0 segundos, la misma mediana de 100.0 segundos y la misma moda de 100 segundos*. Esas medidas de tendencia central no "ven" la diferencia en la variación.

Para mantener nuestras reglas de redondeo tan consistentes y tan simples como sea posible, redondearemos las medidas de variación usando la siguiente regla:

REGLA DE REDONDEO PARA MEDIDAS DE VARIACION. Al redondear el valor de una medida de variación, conserve un decimal más que los presentes en el conjunto original de datos.

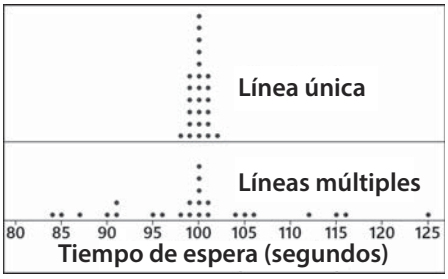


FIGURA 3-2 Gráficas de puntos para los tiempos de espera (segundos) con una línea individual y líneas múltiples

Mejora de la calidad Esta ilustración de una línea única y líneas múltiples es maravillosa porque los bancos realmente cambiaron de múltiples líneas a una sola línea no porque los hiciera más eficientes, no porque los tiempos de espera de los clientes se redujeran, sino porque los clientes son más felices con tiempos de espera con *menos variación*. El cambio no afectó las medidas de tendencia central, pero los bancos instituyeron el cambio para reducir la variación. Un objetivo importante de los negocios y la industria es el siguiente: ***Mejorar la calidad al reducir la variación***.

¿Tienes un segundo?

La unidad de tiempo de un segundo ahora se define como "la duración de 9,192,631,770 periodos de la radiación correspondiente a la transición entre dos niveles hiperfinos del estado estable del átomo de cesio-133". Esta es la primera vez que un segundo se define mediante el comportamiento de los átomos, y no con base en el movimiento de la Tierra; esto da como resultado una exactitud de ±1 segundo en 10,000,000 de años, la unidad de medida más precisa en uso hasta ahora. Debido a su gran exactitud, la definición de un segundo se está utilizando para definir otras cantidades, como el metro. Antes el metro se definía como 1/10,000,000 de la distancia, a lo largo de la superficie de la Tierra, entre el Polo Norte y el Ecuador (pasando por París). En la actualidad, el metro se define como la longitud de la distancia que recorre la luz en el vacío durante un intervalo de tiempo de 1/299,792,458 segundos.

En lo que respecta a los aparatos para medir el tiempo, se ha visto que la desviación estándar tradicional es inaceptable debido a la característica de una tendencia que consiste en una media que cambia con el paso del tiempo. Por ello, se usan otras medidas especiales de la variación, como la varianza Allan, la varianza total y TheoH.

Sin relación con la estadística, pero no menos interesante es el hecho de que los anuncios de relojes suelen mostrar un reloj con una hora cercana a las 10:10. Ese tiempo permite que la marca sea visible, y crea una imagen subliminal de una cara feliz. La hora de las 10:10 ha sido el estándar de la industria desde la década de 1940.

Rango

Comenzaremos con el rango porque es rápido y fácil de calcular, pero no es tan importante como otras medidas de variación.

DEFINICIÓN

El **rango** de un conjunto de valores de datos es la diferencia entre el valor máximo de datos y el valor mínimo de datos.

Rango = (valor máximo de datos) − (valor mínimo de datos)

Propiedad importante del rango

- El rango utiliza sólo los valores máximo y mínimo de los datos, por lo que es muy sensible a los valores extremos. El rango no es *resistente*.
- Debido a que el rango utiliza sólo los valores máximo y mínimo, no toma en cuenta todos los valores y, por lo tanto, no refleja realmente la variación entre todos los valores de los datos.

EJEMPLO 1 Rango

Encuentre el rango de estas velocidades de datos (en Mbps) para Verizon: 38.5, 55.6, 22.4, 14.1, 23.1. (Estas son las primeras cinco velocidades de datos de Verizon listadas en el conjunto de datos 32 "Velocidad de datos en aeropuertos" del apéndice B).

SOLUCIÓN

El rango se encuentra restando el valor más bajo del valor más grande, así que se obtiene

$$\text{Rango} = (\text{valor máximo}) - (\text{valor mínimo}) = 55.6 - 14.1 = 41.50 \text{ Mbps}$$

El rango de 41.50 Mbps se muestra con un decimal más que los presentes en los valores de datos originales.

SU TURNO Encuentre el rango en el ejercicio 5 "Números de jugadores de fútbol americano".

Desviación estándar de una muestra

La *desviación estándar* es la medida de variación más comúnmente utilizada en estadística.

DEFINICIÓN

La **desviación estándar** de un conjunto de valores muestrales, expresada por s, es una medida de cuánto se desvían los valores de datos de la media. Se calcula utilizando la fórmula 3-4 o 3-5. La fórmula 3-5 es solamente una versión diferente de la fórmula 3-4; ambas son algebraicamente iguales.

La desviación estándar encontrada a partir de datos muestrales es un dato estadístico expresado por s, y la desviación estándar determinada a partir de los datos de la población es un parámetro que se expresa con σ. La fórmula de σ es ligeramente diferente porque en la división se usa el tamaño de población N en lugar de $n - 1$. La desviación estándar σ de la población se estudiará más adelante.

Notación

s = desviación estándar *muestral*

σ = desviación estándar *poblacional*

FÓRMULA 3-4

$$s = \sqrt{\frac{\Sigma (x - \bar{x})^2}{n - 1}}$$ desviación estándar muestral

FÓRMULA 3-5

$$s = \sqrt{\frac{n(\Sigma x^2) - (\Sigma x)^2}{n(n - 1)}}$$ fórmula modificada para la desviación estándar muestral (utilizada por calculadoras y software)

Más adelante explicaremos el razonamiento detrás de estas fórmulas, pero por ahora recomendamos el uso de la fórmula 3-4 para un ejemplo o dos, y después aprender a encontrar los valores de desviación estándar usando una calculadora o software.

Propiedades importantes de la desviación estándar

- La desviación estándar es una medida de cuánto se desvían los valores de datos de la *media*.
- El valor de la desviación estándar s nunca es negativo. Es cero sólo cuando todos los valores de datos son exactamente iguales.
- Los mayores valores de s indican mayores cantidades de variación.
- La desviación estándar s puede aumentar dramáticamente con uno o más valores atípicos.
- Las unidades de la desviación estándar s (como minutos, pies, libras) son las mismas que las unidades de los valores de datos originales.
- La desviación estándar muestral s es un **estimador sesgado** de la desviación estándar σ de la población, lo que significa que los valores de la desviación estándar muestral s no se centran en torno al valor de σ. (Esto se explica en la parte 2).

El ejemplo 2 ilustra un cálculo usando la fórmula 3-4 porque esta expresión señala de mejor manera que la desviación estándar se basa en las desviaciones de los valores muestrales con respecto a la media.

Más acciones, menos riesgo

En su libro *Investments*, los autores Zvi Bodie, Alex Kane y Alan Marcus afirman que "la desviación estándar promedio de los rendimientos de carteras compuestas por un solo tipo de acciones fue de 0.554. El riesgo promedio disminuye rápidamente cuando aumenta el número de acciones incluidas en la cartera". También señalan que, con 32 acciones, la desviación estándar es de 0.325, lo que indica mucho menos variación y riesgo. Los autores destacan que con sólo unas cuantas acciones, una cartera tiene alto grado de riesgo "específico de una empresa", lo que significa que el riesgo puede atribuirse a la escasa cantidad de acciones implicadas. Con más de 30 acciones hay muy poco riesgo específico asociado con una sola empresa; en tal situación, casi todo el riesgo es "riesgo de mercado", atribuible al mercado global de acciones. Además, señalan que estos principios son "sólo una aplicación de la bien conocida ley de promedios".

EJEMPLO 2 **Cálculo de la desviación estándar con la fórmula 3-4**

Utilice la fórmula 3-4 para encontrar la desviación estándar de estas velocidades de datos (en Mbps) para Verizon: 38.5, 55.6, 22.4, 14.1, 23.1.

SOLUCIÓN

La columna izquierda de la tabla 3-3 resume el procedimiento general para encontrar la desviación estándar utilizando fórmula 3-4, y la columna derecha ilustra ese procedimiento para los valores muestrales 38.5, 55.6, 22.4, 14.1, 23.1. El resultado que se muestra en la tabla 3-3 es de 16.45 Mbps, que se redondea a un decimal más que los presentes en la lista original de valores muestrales. Además, las unidades para la desviación estándar son las mismas que las unidades de los datos originales. Debido a que los datos originales tienen unidades de Mbps, la desviación estándar es de 16.45 Mbps.

SU TURNO Encuentre la desviación estándar en el ejercicio 5 "Números de jugadores de fútbol americano"

¿Dónde están los bateadores de 0.400?

El último beisbolista que bateó más de 0.400 fue Ted Williams, quien promedió 0.406 en 1941. Hubo promedios por arriba de 0.400 en 1876, 1879, 1887, 1894, 1895, 1896, 1897, 1899, 1901, 1911, 1920, 1922, 1924, 1925 y 1930, pero ninguno desde 1941. ¿Será que ya no existen grandes bateadores? El fallecido Stephen Jay Gould, de la Universidad de Harvard, señaló que el promedio de bateo medio se mantuvo estable en 0.260 durante aproximadamente 100 años, pero la desviación estándar disminuyó de 0.049 en la década de 1870 hasta 0.031 en la actualidad. Él argumentaba que las estrellas de hoy son tan buenas como las del pasado, pero que los mejores lanzadores actuales mantienen los promedios de bateo por debajo de 0.400.

TABLA 3-3

Procedimiento general para determinar la desviación estándar con la fórmula 3-4	Ejemplo específico usando los siguientes valores muestrales: 38.5, 55.6. 22.4. 14.1, 23.1
Paso 1: Calcule la media $\overline{x}$.	La suma de 38.5, 55.6, 22.4, 14.1, 23.1 es 153.7; por lo tanto: $\overline{x} = \frac{\Sigma x}{n} = \frac{38.5 + 55.6 + 22.4 + 14.1 + 23.1}{5} = \frac{153.7}{5} = 30.74$
Paso 2: Reste la media de cada valor muestral individual. [El resultado es una lista de desviaciones de la forma $(x - \overline{x})$.]	Reste la media de 30.74 de cada valor muestral para obtener estas desviaciones respecto a la media: 7.76, 24.86, −8.34, −16.64, −7.64
Paso 3: Eleve al cuadrado cada una de las desviaciones obtenidas en el paso 2. [Esto produce números de la forma $(x - \overline{x})^2$.]	Los cuadrados de las desviaciones del paso 2 son: 60.2176, 618.0196, 69.5556, 276.8896, 58.3696.
Paso 4: Sume todos los cuadrados obtenidos en el paso 3. El resultado es $\Sigma(x - \overline{x})^2$.	La suma de los cuadrados del paso 3 es 1083.0520.
Paso 5: Divida el total del paso 4 por el número $n - 1$, que es 1 menos que el número total de valores muestrales presentes.	Con $n = 5$ valores de datos, $n - 1 = 4$, por lo que dividimos 1083.0520 entre 4 para obtener: $\frac{1083.0520}{4} = 270.7630$.
Paso 6: Encuentre la raíz cuadrada del resultado del paso 5. El resultado es la desviación estándar, expresada por s.	La desviación estándar es $\sqrt{270.7630} = 16.4548777$. Al redondear el resultado, obtenemos $s = 16.45$ Mbps.

EJEMPLO 3 Cálculo de la desviación estándar con la fórmula 3-5

Utilice la fórmula 3-5 para encontrar la desviación estándar de las velocidades de datos (Mbps) para Verizon en el ejemplo 1 (38.5, 55.6, 22.4, 14.1, 23.1).

SOLUCIÓN

Los componentes necesarios en la fórmula 3-5 son:

$n = 5$ (porque hay 5 valores muestrales)

$\Sigma x = 153.7$ (obtenida al sumar los valores originales de la muestra)

$\Sigma x^2 = 5807.79$ (obtenida al sumar los cuadrados de los valores muestrales, $38.5^2 + 55.6^2 + 22.4^2 + 14.1^2 + 23.1^2 = 5807.79$)

Con la fórmula 3-5, obtenemos

$$s = \sqrt{\frac{n(\Sigma x^2) - (\Sigma x)^2}{n(n-1)}} = \sqrt{\frac{5(5807.79) - (153.7)^2}{5(5-1)}} = \sqrt{\frac{5415.26}{20}} = 16.45 \text{ Mbps}$$

El resultado $s = 16.45$ Mbps es el mismo que el del ejemplo 2.

SU TURNO Encuentre la desviación estándar en el ejercicio 5 "Números de jugadores de fútbol americano".

Regla práctica del rango para entender la desviación estándar

La ***regla práctica del rango*** es una herramienta básica pero simple para entender e interpretar la desviación estándar. Se basa en el principio de que, para muchos conjuntos de datos, la gran mayoría (alrededor de 95%) de los valores muestrales se encuentran dentro de 2 desviaciones estándar de la media. Podríamos mejorar la precisión de esta regla teniendo en cuenta factores como el tamaño de la muestra y la distribución, pero aquí sacrificamos la precisión en busca de simplicidad. El concepto de *significancia* que se presenta enseguida será mejorado en capítulos posteriores, especialmente aquellos que incluyen el tema de las pruebas

de hipótesis, que también se llaman pruebas de significancia. La siguiente regla práctica del rango se basa en la media poblacional μ y en la desviación estándar poblacional σ, pero para muestras grandes y representativas, podríamos usar $\bar{x}$ y s en su lugar.

Regla práctica del rango para identificar valores significativos

Los valores ***significativamente bajos*** son $\mu - 2\sigma$ o inferiores.

Los valores **significativamente altos** son $\mu + 2\sigma$ o superiores.

Los **valores no significativos** están entre $(\mu - 2\sigma)$ y $(\mu + 2\sigma)$.

Vea la figura 3-3, que ilustra los criterios anteriores.

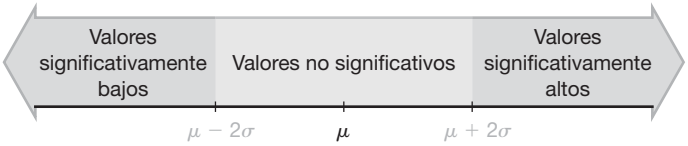


FIGURA 3-3 Regla práctica del rango para identificar los valores significativos

Regla práctica del rango para estimar un valor de la desviación estándar *s*
Para estimar en forma aproximada la desviación estándar de una colección de datos muestrales conocidos, utilice

$$s \approx \frac{\text{rango}}{4}$$

EJEMPLO 4 Regla práctica del rango para interpretar *s*

Si se consideran las 50 velocidades de datos de Verizon listadas en el conjunto de datos 32 "Velocidades de datos en aeropuertos" del apéndice B, la media es 17.60 Mbps y la desviación estándar es 16.02 Mbps. Utilice la regla práctica del rango para encontrar los límites que separan los valores significativamente bajos o significativamente altos; después determine si la velocidad de datos de 77.8 Mbps es significativamente alta.

SOLUCIÓN

Con una media de 17.60 y una desviación estándar de 16.02, utilizamos la regla práctica del rango para encontrar los límites que separan los valores significativamente bajos o significativamente altos, de la siguiente manera:

Los ***valores significativamente bajos*** son $(17.60 - 2 \times 16.02)$ o inferiores.

Por lo tanto, los valores significativamente bajos son -14.44 Mbps o menores.

Los ***valores significativamente altos*** son $(17.60 + 2 \times 16.02)$ o superiores.

Por lo tanto, los valores significativamente altos son 49.64 o mayores.

Los ***valores no significativos*** están entre -14.44 Mbps y 49.64 Mbps.

INTERPRETACIÓN

Con base en estos resultados, esperamos que las velocidades de datos típicas en aeropuertos para Verizon estén entre –14.44 Mbps y 49.64 Mbps. Debido a que el valor dado de 77.8 Mbps queda por encima de 49.64 Mbps, podemos considerarlo significativamente alto.

SU TURNO Resuelva el ejercicio 33 "Pulsos de mujeres".

Variación en las caras

Los investigadores han comentado que "si todos luciéramos más o menos igual, habría un caos total". Se han estudiado las mediciones del cuerpo humano y se ha encontrado que los rasgos faciales varían más que los demás rasgos corporales y la mayor variación ocurre dentro del triángulo formado por los ojos y la boca. Los investigadores afirman que nuestra variación facial desempeña un papel importante en la evolución humana (vea "Morphological and Population Genomic Evidence That Human Faces Have Evolved to Signal Individual Identity", de Sheehan y Nachman, *Nature Communications*, vol. 5, núm. 4800).

EJEMPLO 5 **Regla práctica del rango para estimar *s***

Utilice la regla práctica del rango para estimar la desviación estándar de la muestra de 50 velocidades de datos para Verizon listadas en el conjunto de datos 32 "Velocidades de datos en aeropuertos" del apéndice B. Estos 50 valores tienen un mínimo de 0.8 Mbps y un máximo de 77.8 Mbps.

SOLUCIÓN

La regla práctica del rango indica que podemos estimar la desviación estándar encontrando el rango y dividiéndolo por 4. Con un mínimo de 0.8 y un máximo de 77.8, la desviación estándar s puede estimarse como sigue:

$$s \approx \frac{\text{rango}}{4} = \frac{77.8 - 0.8}{4} = 19.25 \text{ Mbps}$$

INTERPRETACIÓN

El valor real de la desviación estándar es $s = 16.02$ Mbps, por lo que la estimación de 19.25 Mbps está cerca del resultado exacto. Debido a que esta estimación se basa sólo en los valores mínimo y máximo, es generalmente una estimación aproximada que podría estar alejada a una distancia considerable.

SU TURNO Resuelva el ejercicio 29 "Estimación de la desviación estándar".

Desviación estándar de una población

La definición de desviación estándar y las fórmulas 3-4 y 3-5 se aplican a la desviación estándar de datos *muestrales*. Para calcular la desviación estándar σ (sigma minúscula) de una *población*, se usa una fórmula ligeramente diferente: en lugar de dividir por $n - 1$, se divide por el tamaño N de la población, como se muestra:

$$\text{Desviación estándar de la población } \sigma = \sqrt{\frac{\Sigma(x - \mu)^2}{N}}$$

Debido a que generalmente tratamos con datos muestrales, usaremos la fórmula 3-4, en la cual dividimos por $n - 1$. Muchas calculadoras dan tanto la desviación estándar muestral como la poblacional, pero usan varias notaciones.

PRECAUCIÓN Cuando utilice una calculadora para encontrar la desviación estándar, identifique la notación utilizada por su calculadora para obtener la desviación estándar *muestral*, no la desviación estándar *poblacional*.

Varianza de una muestra y una población

Hasta ahora, hemos utilizado el término *variación* como una descripción general de la cantidad en que los valores varían entre sí. (Los términos *dispersión* y *esparcimiento* se usan ocasionalmente en vez de *variación*). El término *varianza* tiene un significado específico.

DEFINICIÓN

La **varianza** de un conjunto de valores es una medida de variación igual al cuadrado de la desviación estándar.

- Varianza muestral: s^2 = cuadrado de la desviación estándar s.
- Varianza poblacional: σ^2 = cuadrado de la desviación estándar poblacional σ.

Notación A continuación se presenta un resumen de la notación para la desviación estándar y la varianza:

s = desviación estándar *muestral*

s^2 = varianza *muestral*

σ = desviación estándar *poblacional*

σ^2 = varianza *poblacional*

Nota: Los artículos de revistas e informes profesionales a menudo utilizan SD para la desviación estándar y VAR para la varianza.

Propiedades importantes de la varianza

- Las unidades de la varianza son los *cuadrados* de las unidades de los valores de datos originales. (Si los valores de datos originales están en pies, la varianza tendrá unidades de pies^2; si están en segundos, la varianza tendrá unidades de seg^2).
- El valor de la varianza puede aumentar dramáticamente con la inclusión de valores atípicos. (La varianza no es *resistente*).
- El valor de la varianza nunca es negativo. Es cero sólo cuando todos los valores de datos son el mismo número.
- La varianza muestral s^2 es un **estimador no sesgado** de la varianza poblacional σ^2, como se describe en la parte 2 de esta sección.

La varianza es un estadístico utilizado en algunos métodos estadísticos, pero para nuestros propósitos actuales, la varianza tiene la seria desventaja de usar unidades que son *diferentes de las unidades del conjunto original de datos*. Esto dificulta entender la varianza en cuanto a su relación con el conjunto de datos original. Debido a esta propiedad, es mejor centrarse primero en la desviación estándar al intentar desarrollar una comprensión de la variación.

PARTE 2 Más allá de lo básico de la variación

En la parte 2 nos centramos en dar sentido a la desviación estándar, de modo que no sea un número misterioso desprovisto de significado práctico. Comenzamos abordando preguntas comunes que se relacionan con la desviación estándar.

¿Por qué la desviación estándar está definida por la fórmula 3-4?

Al medir la variación en un conjunto de datos muestrales, tiene sentido comenzar con las cantidades individuales en las que los valores se desvían de la media. Para un valor de datos particular x, la cantidad de desviación es $x - \bar{x}$. Tiene sentido combinar de alguna manera esas desviaciones en un número que pueda servir como una medida de la variación. La suma de las desviaciones no es buena, porque dicha suma siempre será cero. Para obtener un dato estadístico que mida la variación, es necesario evitar la cancelación de números negativos y positivos. Un método es agregar valores absolutos, como en $\Sigma|x - \bar{x}|$. Si encontramos la media de esa suma, obtenemos la **desviación media absoluta** (o DMA), que es la distancia media de los datos con respecto a la media:

$$\text{Desviación media absoluta} = \frac{\Sigma|x - \bar{x}|}{n}$$

¿Por qué no utilizar la desviación absoluta media en lugar de la desviación estándar?

El cálculo de la desviación absoluta media usa valores absolutos, por lo que emplea una operación que no es "algebraica". La utilización de valores absolutos sería simple, pero crearía dificultades algebraicas en los métodos de estadística inferencial que se estudian en capítulos posteriores. La desviación estándar tiene la ventaja de utilizar sólo operaciones algebraicas.

Debido a que se basa en la raíz cuadrada de una suma de cuadrados, la desviación estándar es muy similar a las fórmulas de distancia encontradas en álgebra. Hay muchos casos en que un procedimiento estadístico se basa en una suma de cuadrados similar. En consecuencia, en vez de utilizar valores absolutos, se elevan al *cuadrado* todas las desviaciones $(x - \bar{x})$ para que sean no negativas, y esos cuadrados se utilizan para calcular la desviación estándar.

¿Por qué dividir por $n - 1$? Después de encontrar todos los valores individuales de $(x - \bar{x})^2$ los combinamos encontrando su suma. Luego dividimos por $n - 1$ porque sólo hay $n - 1$ valores que se pueden asignar sin restricción. Con una media dada, podemos usar cualquier número para los primeros $n - 1$ valores, pero el último valor estará determinado automáticamente. Con la división por $n - 1$, las varianzas muestrales s^2 tienden a centrarse alrededor del valor de la varianza poblacional σ^2; con la división por n, las varianzas muestrales s^2 tienden a *subestimar* el valor de la varianza poblacional σ^2.

¿Cómo encontramos el sentido de un valor de desviación estándar? La parte 1 de esta sección incluyó la regla práctica del rango para interpretar un valor conocido de una desviación estándar o estimar el valor de una desviación estándar. (Vea los ejemplos 4 y 5). Otros dos métodos para interpretar la desviación estándar son la regla empírica y el teorema de Chebyshev.

Regla empírica (o 68-95-99.7) para datos con una distribución en forma de campana

Un concepto útil para interpretar el valor de una desviación estándar es la **regla empírica**. Esta regla establece que *para los conjuntos de datos que tienen una distribución aproximadamente en forma de campana*, se aplican las siguientes propiedades. (Vea la figura 3-4).

- Aproximadamente 68% de todos los valores caen dentro de 1 desviación estándar de la media.
- Aproximadamente 95% de todos los valores caen dentro de 2 desviaciones estándar de la media.
- Aproximadamente 99.7% de todos los valores caen dentro de 3 desviaciones estándar de la media.

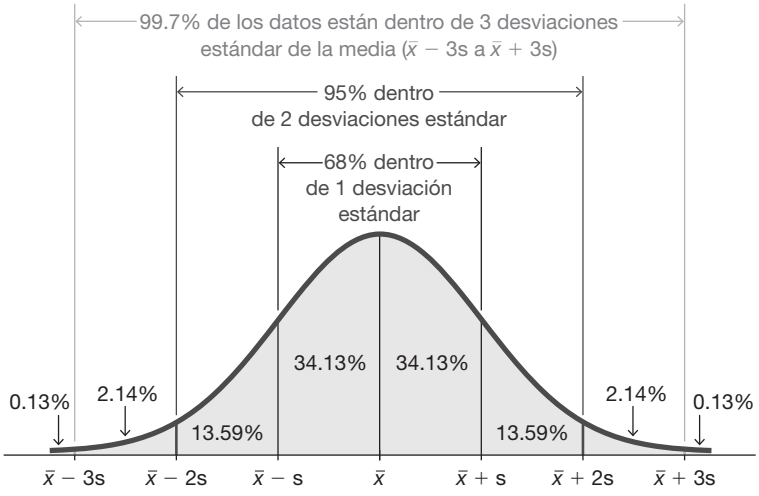


FIGURA 3-4 La regla empírica

EJEMPLO 6 La regla empírica

Las puntuaciones de IQ tienen una distribución en forma de campana con una media de 100 y una desviación estándar de 15. ¿Qué porcentaje de las puntuaciones de IQ está entre 70 y 130?

SOLUCIÓN

La clave para resolver este problema es reconocer que las puntuaciones de 70 y 130 están, cada una, a exactamente dos desviaciones estándar de la media de 100, como se muestra a continuación:

$$2 \text{ desviaciones estándar} = 2s = 2(15) = 30$$

Por lo tanto, 2 desviaciones estándar desde la media es

$$100 - 30 = 70$$
$$\text{o} \quad 100 + 30 = 130$$

La regla empírica nos dice que alrededor de 95% de todos los valores están dentro de 2 desviaciones estándar de la media, por lo que alrededor de 95% de todos los puntajes de IQ están entre 70 y 130.

SU TURNO Resuelva el ejercicio 41 "La regla empírica".

Otro concepto útil para entender o interpretar el valor de una desviación estándar es el **teorema de Chebyshev**. La regla empírica se aplica sólo a conjuntos de datos con distribuciones en forma de campana, pero el teorema de Chebyshev se aplica a *cualquier* conjunto de datos. Desafortunadamente, los resultados del teorema de Chebyshev son sólo aproximados. Debido a que los resultados son límites inferiores ("al menos"), el teorema de Chebyshev tiene una utilidad limitada.

Teorema de Chebyshev

La proporción de cualquier conjunto de datos comprendidos dentro de K desviaciones estándar de la media es siempre al menos $1 - 1/K^2$, donde K es cualquier número positivo mayor que 1. Para $K = 2$ y $K = 3$, obtenemos las siguientes afirmaciones:

- Al menos 3/4 (o 75%) de todos los valores se encuentran dentro de 2 desviaciones estándar de la media.
- Al menos 8/9 (u 89%) de todos los valores se encuentran dentro de 3 desviaciones estándar de la media.

EJEMPLO 7 Teorema de Chebyshev

Las puntuaciones de IQ tienen una media de 100 y una desviación estándar de 15. ¿Qué podemos concluir a partir del teorema de Chebyshev?

SOLUCIÓN

Si se aplica el teorema de Chebyshev con una media de 100 y una desviación estándar de 15, podemos llegar a las siguientes conclusiones:

- Al menos 3/4 (o 75%) de las puntuaciones de IQ están dentro de 2 desviaciones estándar de la media (entre 70 y 130).
- Al menos 8/9 (u 89%) de todos los puntajes de IQ están dentro de 3 desviaciones estándar de la media (entre 55 y 145).

SU TURNO Resuelva el ejercicio 43 "Teorema de Chebyshev".

Comparación de la variación en diferentes muestras o poblaciones

Una buena práctica consiste en comparar dos desviaciones estándar muestrales sólo cuando las medias de la muestra son aproximadamente iguales. Al comparar la variación en muestras o poblaciones con medias muy diferentes, es mejor utilizar el *coeficiente de variación*. Utilice también el coeficiente de variación para comparar la variación de dos muestras o poblaciones con diferentes escalas o unidades de valores, como la comparación de la variación de las *alturas* de los hombres y los *pesos* de los hombres. (Vea el ejemplo 8).

DEFINICIÓN

El **coeficiente de variación** (o ***CV***) para un conjunto de datos muestrales o poblacionales no negativos, expresado como porcentaje, describe la desviación estándar en relación con la media, y está dado por:

Muestra	Población
$CV = \frac{s}{\bar{x}} \cdot 100$	$CV = \frac{\sigma}{\mu} \cdot 100$

REGLA DEL REDONDEO PARA EL COEFICIENTE DE VARIACION Redondee el coeficiente de variación a un decimal (como en 25.3%).

EJEMPLO 8 Velocidades de datos de Verizon y magnitudes de terremoto

Compare la variación de las 50 velocidades de datos de Verizon listadas en el conjunto de datos 32 "Velocidades de datos en aeropuertos" del apéndice B y las magnitudes de 600 terremotos en el conjunto de datos 21 "Terremotos" del apéndice B. Para las velocidades de datos de Verizon $\bar{x} = 17.60$ Mbps y $s = 16.02$ Mbps; para las magnitudes de terremotos, $\bar{x} = 2.572$ y $s = 0.651$. Observe que queremos comparar la variación entre *velocidades de datos* con la variación entre *magnitudes de terremotos*.

SOLUCIÓN

Podemos comparar las desviaciones estándar si se utilizan las mismas escalas y unidades, y si las dos medias son aproximadamente iguales; pero aquí tenemos diferentes escalas y diferentes unidades de medida, por lo que usaremos los coeficientes de variación:

Velocidades de datos de Verizon: $CV = \frac{s}{\bar{x}} \cdot 100\% = \frac{16.02 \text{ Mbps}}{17.60 \text{ Mbps}} \cdot 100\% = 91.0\%$

Magnitudes de terremotos: $CV = \frac{s}{\bar{x}} \cdot 100\% = \frac{0.651}{2.572} \cdot 100\% = 25.3\%$

Ahora podemos ver que las velocidades de datos de Verizon (con $CV = 91.0\%$) varían considerablemente más que las magnitudes de los terremotos (con $CV = 25.3\%$).

Estimadores sesgados y no sesgados

La desviación estándar muestral s es un **estimador sesgado** de la desviación estándar poblacional σ, lo que significa que los valores de la desviación estándar muestral s *no* tienden a centrarse alrededor del valor de la desviación estándar poblacional σ. Aunque los valores individuales de s podrían ser iguales o superiores a σ, los valores de s generalmente tienden a *subestimar* el valor de σ. Por ejemplo, considere una prueba de IQ diseñada para que la desviación estándar de la población sea 15. Si repite el proceso de selección aleatoria de 100 sujetos, dándoles pruebas de IQ y calculando la desviación estándar de la muestra s en cada caso, las desviaciones estándar que obtendrá tenderán a ser inferiores a 15, que es la desviación estándar de la población. No hay corrección que nos permita fijar el sesgo para todas las distribuciones de datos. Existe una corrección que nos permite fijar el sesgo para las poblaciones normalmente distribuidas, pero rara vez se utiliza porque es demasiado compleja y hace modificaciones relativamente menores.

La varianza muestral s^2 es un **estimador no sesgado** de la varianza poblacional σ^2, lo que significa que los valores de s^2 tienden a centrarse alrededor del valor de σ^2 en lugar de tender sistemáticamente a sobreestimarlo o subestimarlo. Considere una prueba de IQ diseñada para que la varianza de la población sea 225. Si repite el proceso de selección aleatoria de 100 sujetos, dándoles pruebas de IQ y calculando la varianza muestral s^2 en cada caso, las varianzas muestrales que obtendrá tenderán a centrarse alrededor 225, que es la varianza de la población.

Los estimadores sesgados y los estimadores no sesgados se analizarán con mayor detalle en la sección 6-3.

CENTRO DE TECNOLOGÍA

Medidas de variación

Acceda a los complementos de software, videos y conjuntos de datos en www.pearsonenespañol.com/triola

Statdisk, Minitab, StatCrunch, Excel y la Calculadora TI-83/84 Plus pueden utilizarse para realizar los cálculos importantes de esta sección. Use los mismos procedimientos de estadística descriptiva dados al final de la sección 3-1 en la página 91.

3-2 Habilidades básicas y conceptos

Conocimiento estadístico y pensamiento crítico

1. Regla práctica del rango para estimar *s* Los 20 volúmenes cerebrales (en cm^3) del conjunto de datos 8 "IQ y tamaño del cerebro" en el apéndice B varían de un mínimo de 963 cm^3 a un máximo de 1439 cm^3. Utilice la regla práctica del rango para estimar la desviación estándar *s* y compare el resultado con la desviación estándar exacta de 124.9 cm^3.

2. Regla práctica del rango para interpretar *s* Los 20 volúmenes cerebrales (en cm^3) del conjunto de datos 8 "IQ y tamaño del cerebro" en el apéndice B tienen una media de 1126.0 cm^3 y una desviación estándar de 124.9 cm^3. Utilice la regla práctica del rango para identificar los límites que separan los valores significativamente bajos y significativamente altos. Para tales datos, ¿un volumen cerebral de 1440 cm^3 sería significativamente alto?

3. Varianza Los 20 sujetos utilizados en el conjunto de datos 8 "IQ y tamaño del cerebro" en el apéndice B tienen pesos con una desviación estándar de 20.0414 kg ¿Cuál es la varianza de sus pesos? Asegúrese de incluir las unidades apropiadas con el resultado.

4. Símbolos Identifique los símbolos utilizados para cada uno de los siguientes estadísticos: (a) desviación estándar muestral; (b) desviación estándar poblacional; (c) varianza muestral; (d) varianza poblacional.

En los ejercicios 5 a 20, encuentre el* rango, *la* varianza *y la* desviación estándar *para los datos muestrales dados. Incluya las unidades apropiadas (como "minutos") en sus resultados. (Los mismos datos se usaron en la sección 3-1, donde encontramos medidas de tendencia central, aquí determinaremos medidas de variación). Luego responda las preguntas dadas.

5. Números de jugadores de fútbol americano A continuación se listan los números de camiseta de 11 jugadores seleccionados aleatoriamente del equipo de los Halcones marinos de Seattle cuando ganaron el Súper Tazón XLVIII. ¿Qué nos dicen los resultados?

89 91 55 7 20 99 25 81 19 82 60

6. Pesos de jugadores de fútbol americano A continuación se listan los pesos en libras de 11 jugadores seleccionados al azar del equipo de los Halcones marinos de Seattle, cuando ganaron el Súper Tazón XLVIII (los mismos jugadores del ejercicio anterior). ¿Es probable que las medidas de variación sean típicas de todos los jugadores de la NFL?

189 254 235 225 190 305 195 202 190 252 305

7. Valor neto de celebridades A continuación se listan los mayores valores netos (en millones de dólares) de las celebridades. Las celebridades son Tom Cruise, Will Smith, Robert De Niro, Drew Carey, George Clooney, John Travolta, Samuel L. Jackson, Larry King, Demi Moore y Bruce Willis. ¿Son las medidas de variación típicas para todas las celebridades?

250 200 185 165 160 160 150 150 150 150

8. Lo que pasa en Las Vegas... A continuación se listan los precios en dólares por una noche en diferentes hoteles ubicados en Las Vegas Boulevard (the "Strip"). ¿Qué tan útiles son las medidas de variación para alguien que busca una habitación?

212 77 121 104 153 264 195 244

9. Huracanes A continuación se lista la cantidad de huracanes que se produjeron en el Atlántico cada año. Los datos se dan en orden anual, a partir del año 2000. ¿Qué característica importante de los datos no es revelada por ninguna de las medidas de variación?

8 9 8 7 9 15 5 6 8 4 12 7 8 2

10. Chícharos en una vaina Los biólogos han realizado experimentos para determinar si una falta de bióxido de carbono en el suelo afecta a los fenotipos de los chícharos. A continuación se listan los códigos del fenotipo, donde 1 = amarillo liso, 2 = verde liso, 3 = amarillo corrugado y 4 = verde corrugado. ¿Se pueden obtener las medidas de variación para estos valores? ¿Los resultados tienen sentido?

2 1 1 1 1 1 1 4 1 2 2 1 2 3 3 2 3 1 3 1 3 1 3 2 2

11. Precios de televisores A continuación se listan los precios de venta (en dólares) de televisores de 60 pulgadas o más, los cuales fueron calificados como "las mejores compras" por la revista *Consumer Reports*. ¿Es probable que las medidas de variación sean típicas para todos los televisores de 60 pulgadas o más?

1800 1500 1200 1500 1400 1600 1500 950 1600 1150 1500 1750

12. Radiación del teléfono celular A continuación se listan las tasas de absorción de radiación medidas (en W/kg) correspondientes a los siguientes teléfonos celulares: iPhone 5S, BlackBerry Z30, Sanyo Vero, Optimus V, Droid Razr, Nokia N97, Samsung Vibrant, Sony Z750a, Kyocera Kona, LG G2 y Virgin Mobile Supreme. Los datos provienen de la Comisión Federal de Comunicaciones. Si a un ejemplar de cada modelo se le mide la absorción de radiación y los resultados se utilizan para encontrar las medidas de variación, ¿son los resultados típicos de la población de teléfonos celulares que están en uso?

1.18 1.41 1.49 1.04 1.45 0.74 0.89 1.42 1.45 0.51 1.38

13. Cafeína en bebidas A continuación se indican las cantidades medidas de cafeína (mg por 12 onzas de bebida) obtenidas en una lata de cada una de 20 marcas (7-UP, A & W Root Beer, Cherry Coke, ...). ¿Son estos datos estadísticos representativos de la población de todas las latas de las mismas 20 marcas consumidas por los estadounidenses?

0 0 34 34 34 45 41 51 55 36 47 41 0 0 53 54 38 0 41 47

14. Muertes de bomberos A continuación se listan las cantidades de heroicos bomberos que perdieron sus vidas en Estados Unidos cada año, mientras combatían incendios forestales. Las cantidades están ordenadas anualmente, comenzando con el año 2000. ¿Qué característica importante de los datos no es revelada por ninguna de las medidas de variación?

20 18 23 30 20 12 24 9 25 15 8 11 15 34

15. Longitud de los pies A continuación se listan las longitudes en pulgadas de los pies de mujeres del ejército seleccionadas al azar, medidos en el sondeo antropométrico de 1988 (ANSUR, abreviatura en inglés). ¿Son los estadísticos representativos de la población actual de todas las mujeres del ejército?

10.4 9.3 9.1 9.3 10.0 9.4 8.6 9.8 9.9 9.1 9.1

16. Universidades más caras A continuación se listan los costos anuales (en dólares) de matrícula y colegiaturas en las 10 universidades más caras de Estados Unidos para un año reciente (con base en datos de *US News & World Report*). Las universidades listadas en orden son Columbia, Vassar, Trinidad, George Washington, Carnegie Mellon, Wesleyana, Tulane, Bucknell, Oberlin y Union. ¿Qué nos dice esta "lista de las 10 primeras" sobre la variación entre los costos para la población de todas las universidades estadounidenses?

49,138 47,890 47,510 47,343 46,962 46,944 46,930 46,902 46,870 46,785

17. Anillo de diamante A continuación se listan las cantidades en dólares que cuestan los paquetes para propuestas matrimoniales en los diferentes estadios de las Ligas Mayores de Béisbol. Cinco de los equipos no permiten propuestas. ¿Existen valores atípicos, y es probable que tengan mucho efecto sobre las medidas de variación?

39 50 50 50 55 55 75 85 100 115 175 175 200
209 250 250 350 400 450 500 500 500 500 1500 2500

18. Ventas de álbumes de discos LP de vinilo A continuación se listan las ventas anuales de discos de vinilo en Estados Unidos (millones de unidades). Las cantidades de álbumes vendidos se presentan en orden cronológico, y la última entrada representa el año más reciente. ¿Las medidas de variación nos dan alguna información sobre una tendencia cambiante a lo largo del tiempo?

0.3 0.6 0.8 1.1 1.1 1.4 1.4 1.5 1.2 1.3 1.4 1.2 0.9 0.9 1.0 1.9 2.5 2.8 3.9 4.6 6.1

19. Fumadores de California En la Encuesta Entrevista sobre Salud en California, se entrevistó a adultos seleccionados al azar. Una de las preguntas fue acerca de cuántos cigarrillos fuman al día, y los resultados se listan a continuación para 50 encuestados seleccionados aleatoriamente. ¿Qué tan bien reflejan los resultados los hábitos del tabaquismo de los adultos en California?

9 10 10 20 40 50 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0

20. Citas rápidas En un estudio sobre citas rápidas realizado en la Universidad de Columbia, se pidió a las mujeres que evaluaran el atractivo de sus acompañantes masculinos, y a continuación se lista una muestra de los resultados (1 = no atractivo, 10 = extremadamente atractivo). ¿Se pueden utilizar los resultados para describir la variación entre los atractivos de la población de varones adultos?

5 8 3 8 6 10 3 7 9 8 5 5 6 8 8 7 3 5 5 6 8 7 8 8 8 7

En los ejercicios 21 a 24, determine el* coeficiente de variación *para cada una de las dos muestras; luego compare la variación. (Los mismos datos se usaron en la sección 3-1).

21. Presión arterial Una muestra de mediciones de presión arterial se toma del conjunto de datos 1 "Datos corporales" del apéndice B, y los valores (mm Hg) se relacionan de manera que 10 sujetos tienen medidas sistólicas y diastólicas.

Sistólica:	118	128	158	96	156	122	116	136	126	120
Diastólica:	80	76	74	52	90	88	58	64	72	82

22. Robo de parquímetros A continuación se listan los montos (en millones de dólares) recaudados en parquímetros por Brinks y otras empresas en la ciudad de Nueva York durante períodos de tiempo similares. Se utilizó un conjunto de datos más amplio para condenar a cinco empleados de Brinks por hurto mayor. Los datos fueron proporcionados por el abogado de la Ciudad de Nueva York, y se pueden encontrar en el sitio web de DASL. ¿Las dos muestras parecen tener diferentes cantidades de variación?

El contratista de recaudación fue Brinks	1.3	1.5	1.3	1.5	1.4	1.7	1.8	1.7	1.7	1.6
El contratista de recaudación no fue Brinks	2.2	1.9	1.5	1.6	1.5	1.7	1.9	1.6	1.6	1.8

23. Pulsos A continuación se listan los pulsos (latidos por minuto) de muestras de hombres y mujeres adultas (del conjunto de datos 1 "Datos corporales" en el apéndice B). ¿Parece haber una diferencia?

Hombres:	86	72	64	72	72	54	66	56	80	72	64	64	96	58	66
Mujeres:	64	84	82	70	74	86	90	88	90	90	94	68	90	82	80

24. Filas en el banco Los tiempos de espera (en segundos) de los clientes en el Banco de Ahorro de Madison se registran con dos configuraciones: línea de clientes única; líneas de clientes individuales.

Línea única	390	396	402	408	426	438	444	462	462	462
Líneas individuales	252	324	348	372	402	462	462	510	558	600

Grandes conjuntos de datos del apéndice B. ***En los ejercicios 25 a 28, considere el conjunto de datos indicado en el apéndice B. Use un software o una calculadora para encontrar el*** **rango,** ***la*** **varianza** ***y la*** **desviación estándar.** ***Exprese las respuestas usando unidades apropiadas, como "minutos".***

25. Tornados Utilice las medidas en la escala F de tornados listadas en el conjunto de datos 22 "Tornados" del apéndice B. Tenga cuidado de considerar los datos faltantes.

26. Terremotos Utilice las magnitudes (en la escala de Richter) de los 600 terremotos listados en el conjunto de datos 21 "Terremotos" del apéndice B. En 1989, el área de la Bahía de San Francisco fue golpeada por un terremoto que midió 7.0 en la escala de Richter. Si añadimos ese valor de 7.0 a los listados en el conjunto de datos, ¿cambian mucho las medidas de variación?

27. Temperaturas corporales Considere el conjunto de datos 3 "Temperaturas corporales" en el apéndice B y use las temperaturas corporales para las 12:00 AM del día 2.

28. Nacimientos Use los pesos al nacer (en gramos) de los 400 bebés listados en el conjunto de datos 4 "Nacimientos" en el apéndice B. Examine la lista de pesos al nacer y haga una observación sobre esos números. ¿Cómo afecta esta observación la forma en que los resultados deben redondearse?

Estimación de la desviación estándar con la regla práctica del rango. ***En los ejercicios 29 a 32, consulte los datos del ejercicio indicado. Después de encontrar el*** **rango** ***de los datos, utilice la regla práctica del rango para calcular el valor de la*** **desviación estándar.** ***Compare el resultado con la desviación estándar calculada con todos los datos.***

29. Ejercicio 25 "Tornados"

30. Ejercicio 26 "Terremotos"

31. Ejercicio 27 "Temperaturas corporales"

32. Ejercicio 28 "Nacimientos"

Identificación de valores significativos con la regla práctica del rango. ***En los ejercicios 33 a 36, use la regla práctica del rango para identificar los límites que separan los*** **valores significativamente bajos** ***y*** **significativamente altos.**

33. Pulso de mujeres Con base en el conjunto de datos 1 "Datos corporales" en el apéndice B, las mujeres tienen pulsos con una media de 74.0 latidos por minuto y una desviación estándar de 12.5 latidos por minuto. ¿Un pulso de 44 latidos por minuto es significativamente bajo o significativamente alto? (Todas estas pulsaciones se miden en reposo).

34. Pulso de hombres Con base en el conjunto de datos 1 "Datos corporales" en el apéndice B, los hombres tienen pulsos con una media de 69.6 latidos por minuto y una desviación estándar de 11.3 latidos por minuto. ¿Un pulso de 50 latidos por minuto es significativamente bajo o significativamente alto? (Todas estas pulsaciones se miden en reposo). Explique.

35. Longitud de los pies Con base en el conjunto de datos 2 "Pies y alturas" en el apéndice B, los varones adultos tienen longitudes de pie con una media de 27.32 cm y una desviación estándar de 1.29 cm. ¿Una longitud de los pies de un adulto de 30 cm es significativamente baja o significativamente alta? Explique.

36. Temperaturas corporales Con base en el conjunto de datos 3 "Temperaturas corporales" en el apéndice B, las temperaturas corporales de los adultos tienen una media de 98.20 °F y una desviación estándar de 0.62 °F. (Se usan los datos de las 12 AM en el día 2). ¿Una temperatura corporal de un adulto de 100 °F es significativamente baja o significativamente alta?

Determinación de la desviación estándar a partir de una distribución de frecuencias. ***En los ejercicios 37 a 40, considere la distribución de frecuencias en el ejercicio dado y encuentre la*** **desviación estándar** ***usando la fórmula siguiente, donde x representa el punto medio de la clase, f es la frecuencia de la clase y n el número total de valores muestrales. Además, compare las desviaciones estándar calculadas con estas desviaciones estándar obtenidas usando la fórmula 3-4 con la lista original de valores de datos: (ejercicio 37) 11.5 años; (ejercicio 38) 8.9 años; (ejercicio 39) 59.5; (Ejercicio 10) 65.4.***

Desviación estándar para la distribución de frecuencias

$$s = \sqrt{\frac{n[\,\Sigma(f \cdot x^2)\,] - [\,\Sigma(f \cdot x)\,]^2}{n(n-1)}}$$

37.

Edad (en años) de la mejor actriz al ganar el Oscar	Frecuencia
20-29	29
30-39	34
40-49	14
50-59	3
60-69	5
70-79	1
80-89	1

38.

Edad (en años) del mejor actor al ganar el Oscar	Frecuencia
20-29	1
30-39	28
40-49	36
50-59	15
60-69	6
70-79	1

39.

Conteo de plaquetas en la sangre de hombres (1000 células/μl)	Frecuencia
0-99	1
100-199	51
200-299	90
300-399	10
400-499	0
500-599	0
600-699	1

40.

Conteo de plaquetas en la sangre de mujeres (1000 células/μl)	Frecuencia
100-199	25
200-299	92
300-399	28
400-499	0
500-599	2

41. La regla empírica Con base en el conjunto de datos 1 "Datos corporales" en el apéndice B, los conteos de plaquetas en la sangre de mujeres tienen una distribución en forma de campana con una media de 255.1 y una desviación estándar de 65.4. (Todas las unidades se dan en 1000 células/μL). Use la regla empírica para determinar cuál es el porcentaje aproximado de mujeres con conteo de plaquetas

a. dentro de 2 desviaciones estándar de la media, o entre 124.3 y 385.9.

b. entre 189.7 y 320.5.

42. La regla empírica Con base en el conjunto de datos 3 "Temperaturas corporales" del apéndice B, las temperaturas corporales de adultos sanos tienen una distribución en forma de campana con una media de 98.20 °F y una desviación estándar de 0.62 °F. Use la regla empírica para determinar cuál es el porcentaje aproximado de adultos sanos con temperatura corporal

a. dentro de 1 desviación estándar de la media, o entre 97.58 °F y 98.82 °F.

b. entre 96.34 °F y 100.06 °F.

43. Teorema de Chebyshev Con base en el conjunto de datos 1 "Datos corporales" en el apéndice B, los conteos de plaquetas en la sangre de mujeres tienen una distribución en forma de campana con una media de 255.1 y una desviación estándar de 65.4. Utilice el teorema de Chebyshev para determinar qué sabemos sobre el porcentaje de mujeres con conteo de plaquetas dentro de 3 desviaciones estándar de la media. ¿Cuáles son los conteos de plaquetas mínimo y máximo que están dentro de 3 desviaciones estándar de la media?

44. Teorema de Chebyshev Con base en el conjunto de datos 3 "Temperaturas corporales" del apéndice B, las temperaturas corporales de adultos sanos tienen una distribución en forma de campana con una media de 98.20 °F y una desviación estándar de 0.62 °F (si se consideran los datos de las 12 AM en el día 2). Utilice el teorema de Chebyshev para determinar qué sabemos sobre el porcentaje de adultos sanos con temperaturas corporales que están dentro de 2 desviaciones estándar de la media. ¿Cuáles son las temperaturas corporales mínima y máxima que están dentro de 2 desviaciones estándar de la media?

3-2 Más allá de lo básico

45. ¿Por qué dividir por $n - 1$? Considere que una población consiste en los valores 9 cigarrillos, 10 cigarrillos y 20 cigarrillos fumados en un día (con base en los datos de la encuesta Entrevista de Salud en California). Suponga que se seleccionan al azar muestras de dos valores con reemplazo de esta población. (Es decir, se reemplaza un valor seleccionado antes de realizar la segunda selección).

a. Encuentre la varianza σ^2 de la población {9 cigarrillos, 10 cigarrillos, 20 cigarrillos}.

b. Después de listar las nueve posibles muestras posibles de dos valores seleccionados con reemplazo, encuentre la varianza muestral s^2 (que incluye la división por $n - 1$) para cada una de ellas; luego encuentre la media de las nueve varianzas muestrales s^2.

c. Para cada una de las nueve diferentes muestras posibles de dos valores seleccionados con reemplazo, encuentre la varianza tratando cada muestra como si fuera una población (usando la fórmula para la varianza poblacional, que incluye la división por n); después determine la media de esas nueve varianzas poblacionales.

d. ¿Qué enfoque resulta en valores que son mejores estimaciones de σ^2: el del inciso (b) o el del inciso (c)? ¿Por qué? Al calcular las varianzas muestrales, ¿debería usarse la división por n o por $n - 1$?

e. Los incisos anteriores muestran que s^2 es un estimador no sesgado de σ^2. ¿Es s un estimador no sesgado de σ? Explique.

46. Desviación media absoluta Utilice la misma población de {9 cigarrillos, 10 cigarrillos, 20 cigarrillos} del ejercicio 45. Muestre que cuando se seleccionan al azar muestras de tamaño 2 con reemplazo, las muestras tienen desviaciones medias absolutas que no se centran alrededor del valor de la desviación media absoluta de la población. ¿Qué indica esto acerca de una desviación media absoluta muestral como estimador de la desviación media absoluta de una población?

3-3 Medidas de posición relativa y gráficas de caja

Concepto clave Esta sección presenta las medidas de posición relativa, que son números que indican la ubicación de los valores de datos en relación con los demás valores dentro del mismo conjunto de datos. El concepto más importante en esta sección es la *puntuación z*, que se utilizará a menudo en los siguientes capítulos. También estudiaremos los cuartiles y percentiles, que son estadísticos comunes, así como un nuevo tipo de gráfica estadística llamada gráfica de caja.

PARTE 1 Fundamentos de las puntuaciones *z*, percentiles, cuartiles y gráficas de caja

Puntuaciones *z*

Una puntuación z se encuentra al convertir un valor a una escala estandarizada, como se establece en la siguiente definición. Esta definición establece que una puntuación z corresponde al número de desviaciones estándar que separan a un dato de la media. Utilizaremos ampliamente las puntuaciones z en el capítulo 6 y en capítulos posteriores.

DEFINICIÓN

Una **puntuación *z*** (o **puntuación estándar** o **valor estandarizado**) es el número de desviaciones estándar que un valor dado *x* se encuentra por arriba o por debajo de la media. Se calcula utilizando las siguientes expresiones:

Muestra		Población
$z = \frac{x - \bar{x}}{s}$	o	$z = \frac{x - \mu}{\sigma}$

REGLA DE REDONDEO PARA LAS PUNTUACIONES *z* Redondee *z* a dos posiciones decimales (como en 2.31)

Esta regla de redondeo está motivada por el formato de tablas estándar en las que los puntajes se expresan con dos decimales, como en la tabla A-2 del apéndice A. El ejemplo 1 ilustra cómo se pueden usar las puntuaciones *z* para comparar valores, incluso si proceden de diferentes poblaciones.

Propiedades importantes de las puntuaciones *z*

1. Una puntuación *z* es el número de desviaciones estándar que un valor dado *x* está por arriba o por debajo de la media.
2. Las puntuaciones *z* se expresan como números sin unidades de medida.
3. Un valor de datos es *significativamente bajo*, si su puntuación *z* es menor o igual a -2 o el valor es significativamente alto si su puntuación *z* es mayor o igual a $+2$.
4. Si un valor de datos individual es menor que la media, su puntuación correspondiente es un número negativo.

EJEMPLO 1 Comparación del peso de un bebé y la temperatura corporal de un adulto

¿Cuál de los siguientes dos valores de datos es más extremo en relación con el conjunto de datos del que procede?

- El peso de 4000 g de un bebé recién nacido (entre 400 pesos con media muestral $\bar{x} = 3152.0$ g y desviación estándar muestral $s = 693.4$ g)
- La temperatura de 99 °F de un adulto (entre 106 adultos con media muestral $\bar{x} = 98.20$ °F y desviación estándar muestral $s = 0.62$ °F)

SOLUCIÓN

El peso de 4000 g y la temperatura corporal de 99 °F pueden estandarizarse convirtiendo cada uno de ellos a puntuaciones *z* como se muestra a continuación.

4000 g de peso al nacer:

$$z = \frac{x - \bar{x}}{s} = \frac{4000 \text{ g} - 3152.0 \text{ g}}{693.4 \text{ g}} = 1.22$$

99 °F de temperatura corporal:

$$z = \frac{x - \bar{x}}{s} = \frac{99 \text{ °F} - 98.20 \text{ °F}}{0.62 \text{ °F}} = 1.29$$

INTERPRETACIÓN

Las puntuaciones z muestran que el peso al nacer de 4000 g está 1.22 desviaciones estándar por arriba de la media, y la temperatura corporal de 99 °F está 1.29 desviaciones estándar por arriba de la media. Debido a que la temperatura del cuerpo está más alejada de la media, es el valor más extremo. Una temperatura corporal de 99 °F es ligeramente más extrema que un peso al nacer de 4000 g.

SU TURNO Resuelva el ejercicio 13 "Los hombres más altos y más bajos".

Uso de puntuaciones z para identificar valores significativos En la sección 3-2 usamos la regla práctica del rango para concluir que un valor es significativamente bajo o significativamente alto si está por lo menos 2 desviaciones estándar alejado de la media. Se deduce que los valores significativamente bajos tienen puntuaciones z menores o iguales a -2 y que los valores significativamente altos tienen puntuaciones z mayores o iguales a $+2$, como se ilustra en la figura 3-5. Si se usa este criterio con los dos valores individuales usados en el ejemplo 1, se observa que ninguno de los dos valores es significativo porque ambas puntuaciones z están entre -2 y $+2$.

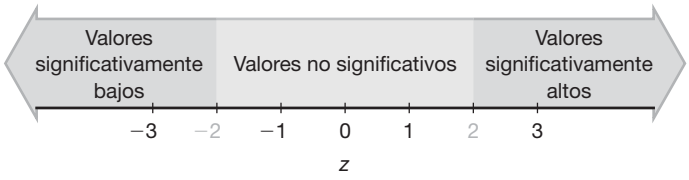


FIGURA 3-5 Interpretación de puntuaciones z

Los valores significativos son aquellos con puntuaciones $z \leq -2.00$ o ≥ 2.00.

EJEMPLO 2 ¿Un conteo de plaquetas de 75 es significativamente bajo?

El conteo de plaquetas más bajo en el conjunto de datos 1 "Datos corporales" en el apéndice B es 75. (El conteo de plaquetas se mide en 1000 células/μl). ¿Es ese valor significativamente bajo? Con base en los conteos de plaquetas del conjunto de datos 1 del apéndice B, suponga que los conteos de plaquetas tienen una media de $\bar{x} = 239.4$ y una desviación estándar de $s = 64.2$.

SOLUCIÓN

El conteo de plaquetas de 75 se convierte en una puntuación z como se muestra a continuación:

$$z = \frac{x - \bar{x}}{s} = \frac{75 - 239.4}{64.2} = -2.56$$

INTERPRETACIÓN

El conteo de plaquetas de 75 se convierte en la puntuación z de -2.56. Consulte la figura 3-5 para ver que $z = -2.56$ es menor que -2, por lo que el conteo de plaquetas de 75 es significativamente bajo. (El conteo bajo de plaquetas se llama trombocitopenia, no por falta de un término mejor).

SU TURNO Resuelva el ejercicio 9 "ACT".

Una puntuación z es una medida de posición, en el sentido de que describe la ubicación de un valor (en términos de desviaciones estándar) con relación a la media. Los percentiles y cuartiles son otras medidas de posición útiles para comparar valores dentro del mismo conjunto de datos o entre diferentes conjuntos de datos.

Percentiles

Los percentiles son un tipo de *cuantiles* —o *fractiles*— que dividen los datos en grupos con aproximadamente el mismo número de valores en cada grupo.

DEFINICIÓN

Los **percentiles** son medidas de ubicación, expresadas como P_1, P_2, . . . , P_{99}, que dividen un conjunto de datos en 100 grupos con aproximadamente el 1% de los valores en cada grupo.

El percentil 50, denominado P_{50}, tiene aproximadamente el 50% de los valores de datos por debajo de él y aproximadamente el 50% de los valores de datos por encima de él, por lo que el percentil 50 es igual que la mediana. No hay un acuerdo universal sobre el procedimiento que debe usarse para calcular los percentiles, pero describiremos procedimientos relativamente simples para (1) encontrar el percentil de un valor de datos y (2) convertir un percentil en su valor de datos correspondiente. Comenzamos con el primer procedimiento.

Determinación del percentil de un valor de datos

El proceso de encontrar el percentil que corresponde a un valor de datos particular x está dado por lo siguiente (redondee el resultado al número entero más próximo):

$$\text{Percentil de valor } x = \frac{\text{número de valores menores que } x}{\text{número total de valores}} \cdot 100$$

EJEMPLO 3 Determinación de un percentil

En la tabla 3-4 se listan las mismas velocidades de datos para teléfonos celulares de Verizon que aparecen en el conjunto de datos 32 "Velocidades de datos en aeropuertos", pero en la tabla 3-4 las velocidades de datos se colocan en orden creciente. Encuentre el percentil para la velocidad de datos de 11.8 Mbps.

TABLA 3-4 Velocidades de datos en aeropuertos para Verizon *ordenadas* (Mbps)

0.8	1.4	1.8	1.9	3.2	3.6	4.5	4.5	4.6	6.2
6.5	7.7	7.9	9.9	10.2	10.3	10.9	11.1	11.1	11.6
11.8	12.0	13.1	13.5	13.7	14.1	14.2	14.7	15.0	15.1
15.5	15.8	16.0	17.5	18.2	20.2	21.1	21.5	22.2	22.4
23.1	24.5	25.7	28.5	34.6	38.5	43.0	55.6	71.3	77.8

SOLUCIÓN

En la lista ordenada de velocidades de datos en aeropuertos de la tabla 3-4, se observa que hay 20 velocidades de datos menores que 11.8 Mbps, así que

$$\text{Percentil de } 11.8 = \frac{20}{50} \cdot 100 = 40$$

INTERPRETACIÓN

Una velocidad de datos de 11.8 Mbps está en el percentil 40. Esto puede interpretarse de manera flexible: una velocidad de datos de 11.8 Mbps separa el 40% de los valores más bajos del 60% de los valores más altos. Se tiene $P_{40} = 11.8$ Mbps.

SU TURNO Resuelva el ejercicio 17 "Percentiles".

El ejemplo 3 muestra cómo convertir de un valor muestral dado al percentil correspondiente. Existen varios métodos diferentes para el procedimiento inverso de convertir un percentil dado al valor correspondiente en el conjunto de datos. El procedimiento que utilizaremos se resume en la figura 3-6, que utiliza la siguiente notación.

Índice del costo de la risa

En realidad hay un Índice del Costo de la Risa (ICR), que busca los costos de artículos como pollos de plástico, anteojos de Groucho Marx, entradas a clubes de comediantes y otros 13 indicadores principales del humor. Este es el mismo método básico que se utiliza en la creación del Índice de Precios al Consumidor (IPC), el cual se basa en un promedio ponderado de bienes y servicios adquiridos por consumidores comunes. Mientras que las puntuaciones estándar y los percentiles nos permiten comparar valores diferentes ignorando cualquier elemento del tiempo, los números índice, como el ICR y el IPC, nos permiten comparar el valor de alguna variable con su valor en un periodo tomado como base. El valor de un número índice es el valor actual dividido entre el valor base, multiplicado por 100.

Notación

n número total de valores en el conjunto de datos

k percentil que se utiliza (ejemplo: para el percentil 25, $k = 25$).

L localizador que da la *posición* de un valor (ejemplo: para el duodécimo valor en la lista ordenada, $L = 12$).

P_k k-ésimo percentil (ejemplo: P_{25} es el percentil 25).

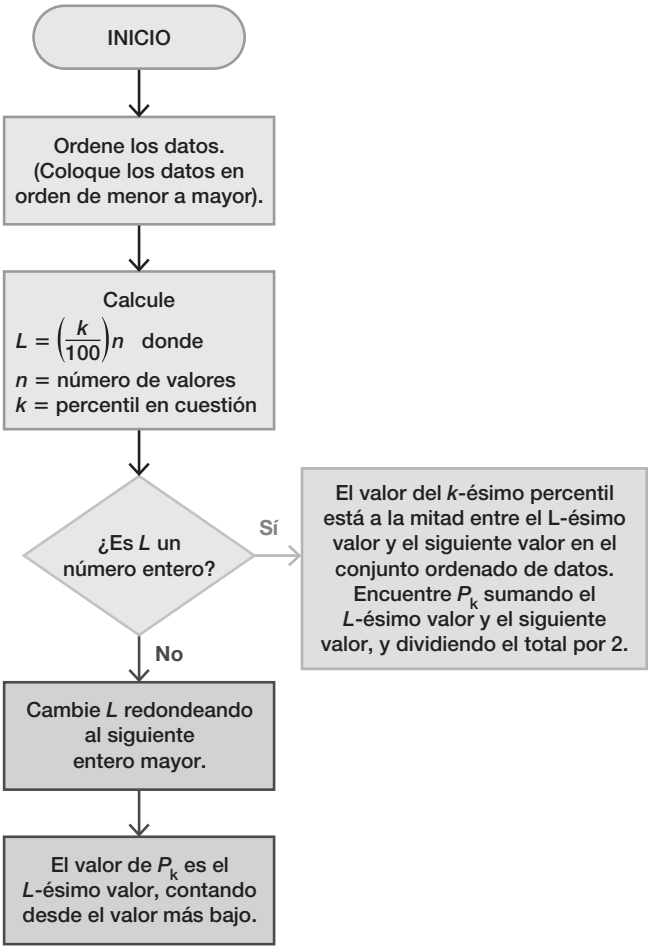


FIGURA 3-6 Conversión del *k*-ésimo percentil al valor de datos correspondiente

EJEMPLO 4 Conversión de un percentil a un valor de datos

Consulte las velocidades de datos ordenadas en la tabla 3-4 y use el procedimiento de la figura 3-6 para encontrar el valor del vigésimo quinto percentil, P_{25}.

SOLUCIÓN

A partir de la figura 3-6, se observa que los datos de la muestra ya están ordenados, por lo que podemos proceder a encontrar el valor del localizador L. En este cálculo usamos $k = 25$ porque estamos tratando de encontrar el valor del percentil 25. Utilizamos $n = 50$ porque hay 50 valores de datos.

$$L = \frac{k}{100} \cdot n = \frac{25}{100} \cdot 50 = 12.5$$

Puesto que $L = 12.5$ no es un número entero, pasamos al siguiente cuadro inferior de la figura 3-6, donde cambiamos L al redondear de 12.5 al siguiente entero mayor: 13. (En este libro redondeamos de la manera habitual, pero éste es uno de los dos casos en los que redondeamos hacia arriba). En el cuadro inferior vemos que el valor de P_{25} es el valor número 13, contando desde el valor más bajo. En la tabla 3-4, el valor 13 es 7.9. Es decir, $P_{25} = 7.9$ Mbps. A grandes rasgos, alrededor de 25% de las velocidades de datos son menores que 7.9 Mbps y 75% de ellas son mayores que 7.9 Mbps.

SU TURNO Resuelva el ejercicio 23 "Cuartil".

EJEMPLO 5 Conversión de un percentil a un valor de datos

Consulte las velocidades de datos ordenadas de la tabla 3-4. Utilice la figura 3-6 para encontrar el percentil 40, expresada por P_{40}.

SOLUCIÓN

Con referencia a la figura 3-6, vemos que los datos muestrales ya están ordenados, por lo que podemos proceder a calcular el valor del localizador L. En este cálculo, usamos $k = 40$ porque estamos tratando de encontrar el valor del percentil 40, y utilizamos $n = 50$ porque hay 50 valores de datos.

$$L = \frac{k}{100} \cdot n = \frac{40}{100} \cdot 50 = 20$$

Puesto que $L = 20$ es un número entero, procedemos a la caja de la figura 3-6 ubicada a la derecha. Ahora vemos que el valor del percentil 40 está a la mitad entre el L-ésimo valor (vigésimo) y el siguiente valor en el conjunto original de datos. Es decir, el valor del percentil 40 está a la mitad entre el valor 20 y el valor 21. El valor 20 en la tabla 3-4 es 11.6 y el valor 21 es 11.8, por lo que el valor a la mitad entre ellos es 11.7 Mbps. Concluimos que el percentil 40 es $P_{40} = 11.7$ Mbps.

SU TURNO Resuelva el ejercicio 21 "Percentil".

Cuartiles

De la misma manera que hay 99 percentiles que dividen los datos en 100 grupos, hay tres cuartiles que dividen los datos en cuatro grupos.

DEFINICIÓN

Los **cuartiles** son medidas de ubicación, denominadas Q_1, Q_2 y Q_3, que dividen un conjunto de datos en cuatro grupos con aproximadamente el 25% de los valores en cada uno.

A continuación se describen los cuartiles de manera más precisa que en la definición anterior:

- **Q_1 (Primer cuartil):** mismo valor que P_{25}. Separa el 25% de los valores inferiores ordenados del 75% superior. (Para ser más precisos, al menos 25% de los valores ordenados son menores o iguales a Q_1, y al menos 75% de los valores son mayores o iguales a Q_1).
- **Q_2 (Segundo cuartil):** igual a P_{50} e igual a la mediana. Separa el 50% de los valores inferiores ordenados del 50% superior.
- **Q_3 (Tercer cuartil):** igual a P_{75}. Separa el 75% de los valores inferiores ordenados del 25% superior. (Para ser más precisos, al menos 75% de los valores ordenados son menores o iguales a Q_3 y al menos 25% de los valores son mayores o iguales a Q_3).

Niveles Nielsen para estudiantes universitarios

Los niveles Nielsen representan una de las medidas más importantes del teleauditorio y afectan a miles de millones de dólares en publicidad televisiva. En el pasado, los hábitos televisivos de los estudiantes universitarios eran ignorados, con el resultado de que un gran segmento de la audiencia joven quedara excluido. Nielsen Media Research está ahora incluyendo a los estudiantes universitarios que no viven en casa.

Algunos programas de televisión tienen gran atractivo para los televidentes en el grupo de edad de 18 a 24 años, y los niveles de audiencia de estos programas han aumentado sustancialmente con la inclusión de los estudiantes universitarios. Para los varones, la difusión del *fútbol americano del domingo por la noche* a través de la NBC tuvo un aumento del 20% después de incluir a los estudiantes universitarios. Los niveles más altos de audiencia se traducen en mayores ganancias por mayores cobros a los patrocinadores comerciales. Estos niveles también dan reconocimiento a los estudiantes universitarios, lo cual afecta la programación que reciben.

$Q_1 = P_{25}$
$Q_2 = P_{50}$
$Q_3 = P_{75}$

La determinación de los valores de los cuartiles se puede lograr con el mismo procedimiento utilizado para encontrar percentiles. Simplemente utilice las relaciones mostradas al margen. En el ejemplo 4 se encontró que $P_{25} = 7.9$ Mbps, por lo que se deduce que $Q_1 = 7.9$ Mbps.

PRECAUCIÓN Así como no hay un acuerdo universal sobre un procedimiento para encontrar percentiles, tampoco existe un solo procedimiento para calcular cuartiles, y las diferentes tecnologías a menudo dan resultados diferentes. Si utiliza una calculadora o un software que incluyan cuartiles, puede obtener resultados que difieran un poco de las respuestas obtenidas utilizando los procedimientos descritos aquí.

En secciones anteriores de este capítulo describimos varios estadísticos, incluyendo la media, la mediana, la moda, el rango y la desviación estándar. Algunos otros estadísticos se definen utilizando cuartiles y percentiles, como los siguientes:

$$\text{Rango intercuartil (o RIQ)} = Q_3 - Q_1$$

$$\text{Rango semi-intercuartil} = \frac{Q_3 - Q_1}{2}$$

$$\text{Cuartil medio} = \frac{Q_3 + Q_1}{2}$$

$$\text{Rango percentil 10-90} = P_{90} - P_{10}$$

Resumen de 5 números y gráfica de caja

Los valores del mínimo, el máximo y los tres cuartiles (Q_1, Q_2, Q_3) se utilizan para el resumen de 5 números y la construcción de gráficas de caja.

DEFINICIÓN

Para un conjunto de datos, el **resumen de 5 números** consta de los siguientes cinco valores:

1. Mínimo
2. Primer cuartil, Q_1
3. Segundo cuartil, Q_2 (igual a la mediana)
4. Tercer cuartil, Q_3
5. Máximo

EJEMPLO 6 Determinación de un resumen de 5 números

Utilice las velocidades de datos en aeropuertos para Verizon de la tabla 3-4 para encontrar el resumen de 5 números.

SOLUCIÓN

Debido a que las velocidades de datos en aeropuertos para Verizon de la tabla 3-4 están ordenadas, es fácil ver que el mínimo es 0.8 Mbps y el máximo es 77.8 Mbps. El valor del primer cuartil es $Q_1 = 7.9$ Mbps (del ejemplo 4). La mediana es igual a Q_2, y es 13.9 Mbps. Además, podemos encontrar que $Q_3 = 21.5$ Mbps usando el mismo procedimiento para encontrar P_{75} (tal como lo indica la figura 3-6). Por lo tanto, el resumen de 5 números es 0.8, 7.9, 13.9, 21.5 y 77.8 (todos en unidades de Mbps).

SU TURNO Encuentre el resumen de 5 números en el ejercicio 29 “Citas rápidas”.

Los valores del resumen de 5 números se utilizan para la construcción de una gráfica de caja, definido como sigue.

DEFINICIÓN

Una **gráfica de caja** (o **diagrama de caja y bigotes**) es una gráfica de un conjunto de datos que consiste en una línea que se extiende desde el valor mínimo hasta el valor máximo, y una caja con líneas dibujadas en el primer cuartil Q_1, la mediana y el tercer cuartil Q_3. (Vea la figura 3-7).

Procedimiento para elaborar una gráfica de caja

1. Encuentre el resumen de 5 números (valor mínimo, Q_1, Q_2, Q_3, valor máximo).
2. Construya un segmento de línea que se extienda desde el valor mínimo hasta el valor máximo de los datos.
3. Construya una caja (rectángulo) que se extienda de Q_1 a Q_3, y dibuje una línea en la caja sobre el valor de Q_2 (la mediana).

PRECAUCIÓN Debido a que no hay un acuerdo universal sobre los procedimientos para encontrar cuartiles, y como las gráficas de caja se basan en cuartiles, los diversos paquetes y aplicaciones de software pueden producir gráficas distintas.

EJEMPLO 7 Elaboración de una gráfica de caja

Utilice las velocidades de datos en aeropuertos para Verizon listadas en la tabla 3-4 para construir una gráfica de caja.

SOLUCIÓN

La gráfica de caja usa el resumen de 5 números encontrado en el ejemplo 6: 0.8, 7.9, 13.9, 21.5 y 77.8 (todos en unidades de Mbps). La figura 3-7 es la gráfica de caja que representa las velocidades de datos en aeropuertos para Verizon listadas en la tabla 3-4.

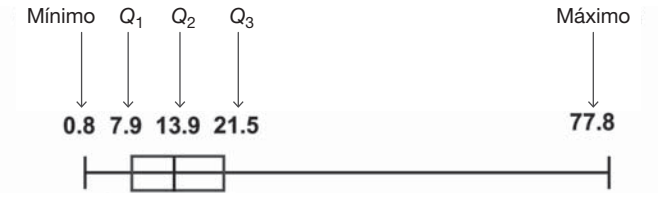


FIGURA 3-7 Gráfica de caja de las velocidades de datos en aeropuertos para Verizon (Mbps)

SU TURNO Construya la gráfica de caja en el ejercicio 29 "Citas rápidas".

Asimetría A menudo, una gráfica de caja se puede usar para identificar asimetrías. Recuerde que en la sección 2-2 establecimos que una distribución de datos es **asimétrica** si se extiende más hacia un lado que hacia el otro. En un histograma de datos asimétrico a la derecha (también denominado *positivamente asimétrico*), hay una cola derecha más larga que indica que relativamente pocos valores de datos son altos; la mayoría de los valores se encuentran a la izquierda. La gráfica de caja de la figura 3-7 muestra que los datos son asimétricos a la derecha y la mayoría de los valores se encuentran a la izquierda.

Debido a que la forma de una gráfica de caja está determinada por los valores del resumen de cinco números, no es una gráfica de la distribución de los datos, y no muestra información tan detallada como un histograma o un diagrama de tallo y hojas. Sin embargo, las gráficas de caja son útiles para comparar dos o más conjuntos de datos. Cuando utilice dos o más gráficas de caja para comparar diferentes conjuntos de datos, grafíquelos en la misma escala para facilitar las comparaciones. Los métodos que se estudian más adelante en este libro permiten analizar las comparaciones de los conjuntos de datos más formalmente que las conclusiones subjetivas basadas en una gráfica. Siempre es aconsejable construir gráficas adecuadas, como histogramas, gráficas de puntos y gráficas de caja, pero no debemos confiar únicamente en juicios subjetivos basados en gráficas.

EJEMPLO 8 Comparación de las velocidades de datos de Verizon, Sprint, AT&T y T-Mobile

El problema del capítulo se refiere a las velocidades de datos para teléfonos inteligentes en 50 aeropuertos, y las velocidades se miden para los operadores Verizon, Sprint, AT&T y T-Mobile. Utilice la misma escala para construir las cuatro gráficas de caja correspondientes; luego compare los resultados.

SOLUCIÓN

Los diagramas de caja generados por Statdisk que se muestran en la figura 3-8 sugieren que las velocidades de datos de Verizon son generalmente más rápidas.

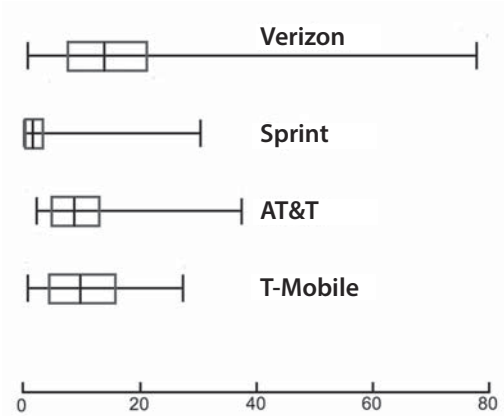


FIGURA 3-8 Gráficas de caja de las velocidades de datos en aeropuertos para diferentes operadores

SU TURNO Resuelva el ejercicio 33 “Pulsos”.

Valores atípicos

Cuando se analizan datos, es importante identificar y considerar los valores atípicos porque pueden afectar significativamente los valores de algunos datos estadísticos importantes (como la media y la desviación estándar) y también pueden tener un gran impacto en los métodos importantes que se estudian más adelante en este libro. En el capítulo 2 describimos los valores atípicos como valores muestrales que están muy alejados de la gran mayoría de los demás valores de un conjunto de datos, pero esa descripción es vaga y no proporciona criterios objetivos específicos. La parte 2 de esta sección incluye una descripción de las *gráficas de caja modificadas* junto con una definición más precisa de los valores atípicos utilizados en el contexto de la creación de este tipo de gráficas.

PRECAUCIÓN Cuando analice datos, siempre identifique los valores atípicos y considere sus efectos, los cuales pueden ser sustanciales.

PARTE 2 Valores atípicos y gráficas de caja modificadas

Hemos observado que la descripción de los valores atípicos es algo imprecisa, pero con la intención de construir *gráficas de caja modificadas*, podemos considerar valores atípicos como valores de datos que cumplen con criterios específicos basados en los cuartiles y el rango intercuartil. (A menudo se denomina RIQ = $Q_3 - Q_1$).

Identificación de valores atípicos para las gráficas de caja modificadas

1. Encuentre los cuartiles Q_1, Q_2 y Q_3.
2. Determine el rango intercuartil (RIQ), donde RIQ = $Q_3 - Q_1$.
3. Evalúe 1.5 × RIQ.
4. **En una gráfica de caja modificada, un valor de datos es *atípico* si está por arriba de Q_3, en una cantidad superior a 1.5 × RIQ o por debajo de Q_1, en una cantidad superior a 1.5 × RIQ**

Gráficas de caja modificadas

Las gráficas descritas anteriormente se denominan **gráficas de caja esqueléticas** (o **regulares**), pero algunos paquetes de software estadístico proporcionan gráficas modificadas, que representan valores extremos como puntos especiales. Una **gráfica de caja modificada** es una gráfica de caja regular construida con las siguientes modificaciones: (1) Un símbolo especial (un asterisco o un punto) se utiliza para identificar los valores atípicos según se definieron anteriormente, y (2) la línea horizontal sólida sólo se extiende hasta el valor de datos mínimo que no es atípico y hasta el valor de datos máximo que tampoco es atípico. (*Nota: los ejercicios que implican gráficas de caja modificadas se encuentran solamente en los ejercicios "Más allá de lo básico"*).

EJEMPLO 9 Construcción de una gráfica de caja modificada

Utilice las velocidades de datos en aeropuertos para Verizon en el conjunto de datos 32 "Velocidades de datos en aeropuertos" del apéndice B para construir una gráfica de caja modificada.

SOLUCIÓN

Comencemos con los cuatro pasos descritos anteriormente para identificar los valores atípicos en una gráfica de caja modificada.

1. A partir de las velocidades de datos para Verizon, los tres cuartiles son $Q_1 = 7.9$, $Q_2 = 13.9$ (mediana) y $Q_3 = 21.5$. (Todos los valores se dan en Mbps y estos cuartiles se encontraron en el ejemplo 6).
2. El rango intercuartil es RIQ = $Q_3 - Q_1 = 21.5 - 7.9 = 13.6$.
3. 1.5 × RIQ = 1.5 × 13.6 = 20.4.
4. Cualesquiera valores atípicos están arriba de $Q_3 = 21.5$ por más de 20.4, o debajo de $Q_1 = 7.9$ por más de 20.4. Esto significa que cualquier valor atípico es mayor que 41.9 o menor que −12.5 (lo que es imposible; entonces, aquí no hay valores atípicos en el extremo inferior).

Ahora es posible examinar las velocidades de datos en aeropuertos para Verizon originales a fin de identificar las velocidades superiores a 41.9, y se encuentra: 43.0, 55.6, 71.3 y 77.8, que son los únicos valores atípicos.

Ahora podemos construir la gráfica de caja modificada que se muestra en la figura 3-9 de la página siguiente. Ahí, los cuatro valores atípicos se identifican como puntos especiales, los tres cuartiles se muestran como en una gráfica regular y la línea horizontal se extiende desde el valor de datos más bajo que no es atípico (0.8) hasta el valor de datos más alto que tampoco es atípico (38.5).

continúa

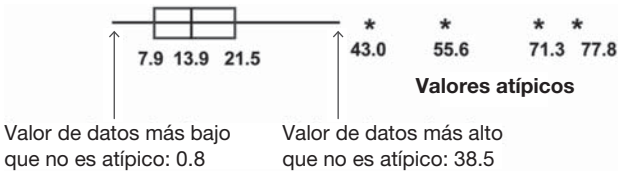


FIGURA 3-9 Gráfica de caja modificada de las velocidades de datos en aeropuertos para Verizon (Mbps)

SU TURNO Resuelva el ejercicio 37 "Valores atípicos y gráficas de caja modificadas".

PRECAUCIÓN Debido a que no existe un acuerdo universal sobre los procedimientos para encontrar cuartiles, y como las gráficas de caja modificadas se basan en cuartiles, las diferentes tecnologías pueden producir gráficas de caja modificadas distintas.

CENTRO DE TECNOLOGÍA

Gráficas de caja, resumen de 5 números, valores atípicos

Acceda a los complementos técnicos, videos y conjuntos de datos en www.pearsonenespañol.com/triola

Statdisk

Resumen de 5 números

Use el procedimiento de **estadística descriptiva** dado al final de la sección 3-1 en la página 91.

Gráficas de caja

1. Haga clic en **Data** del menú superior.
2. Seleccione **Boxplot** en el menú desplegable.
3. Seleccione las columnas de datos deseadas.
4. Haga clic en **Boxplot** o **Modified Boxplot**.

Valores atípicos

Cree una gráfica de caja modificada utilizando el procedimiento anterior u ordénelo de la siguiente manera:

1. Haga clic en **Data** del menú superior.
2. Seleccione **Sort Data** en el menú desplegable.
3. Haga clic en **Sort** después de elegir las opciones deseadas en el menú de ordenación.
4. Examine los valores mínimo y máximo para determinar si están muy alejados de los demás valores.

Minitab

Resumen de 5 números

Use el procedimiento de **estadística descriptiva** dado al final de la sección 3-1 en la página 91.

Gráficas de caja

1. Haga clic en **Graph** del menú superior
2. Seleccione **Boxplot** en el menú desplegable
3. Seleccione la opción **Simple** para una o varias gráficas de caja, luego haga clic en **OK**.
4. Haga doble clic en la(s) columna(s) de datos deseada(s) para que aparezca en la ventana de *Graph variables* y luego haga clic en **OK**.

Valores atípicos

Cree una gráfica de caja modificada utilizando el procedimiento anterior u ordénelo de la siguiente manera:

1. Haga clic en **Data** del menú superior.
2. Seleccione **Sort** en el menú desplegable.
3. Haga doble clic en la columna de datos deseada para que aparezca en la ventana de *Sort Column(s)*.
4. Haga clic en el cuadro *By column* y seleccione la misma columna de datos. Haga clic en **OK**.
5. Examine los valores mínimo y máximo para determinar si están muy alejados de otros valores.

CENTRO DE TECNOLOGÍA *continuación*

Gráficas de caja, resumen de 5 números, valores atípicos

Acceda a los complementos tecnológicos, videos y conjuntos de datos en www.pearsonenespañol.com/triola

StatCrunch

Resumen de 5 números

Use el procedimiento de **estadística descriptiva** dado al final de la sección 3-1 en la página 91.

Gráficas de caja

1. Haga clic en **Graph** del menú superior.
2. Seleccione **Boxplot** en el menú desplegable.
3. Seleccione la columna de datos deseada. Para una gráfica de caja modificada, marque la casilla *Use fences*.
4. Haga clic en **Compute!**

Valores atípicos

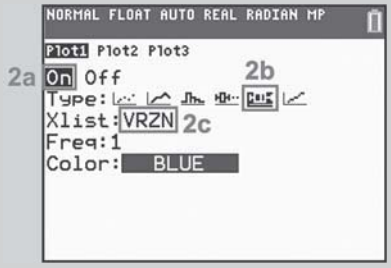
Cree una gráfica de caja modificada utilizando el procedimiento anterior u ordénelo de la siguiente manera:

1. Haga clic en **Data** del menú superior.
2. Seleccione **Sort** en el menú desplegable.
3. Seleccione la columna de datos deseada.
4. Haga clic en **Compute!**
5. Examine los valores mínimo y máximo para determinar si están muy alejados de los demás valores.

Calculadora TI-83/84 Plus

Resumen de 5 números

Use el procedimiento de **estadística descriptiva** dado al final de la sección 3-1 en la página 91.



Gráficas de caja

1. Abra el menú **STAT PLOTS** pulsando 2ND, Y=
2. Presione ENTER para acceder a la pantalla de configuración *Plot 1* como se muestra:
   a. Seleccione **ON** y pulse ENTER.
   b. Seleccione el segundo ícono de gráfica de caja, presione ENTER. Seleccione el primer ícono de gráfica de caja para una gráfica modificada.
   c. Introduzca el nombre de la lista que contiene los datos.
3. Presione ZOOM luego 9 (ZoomStat) para mostrar la gráfica de caja.
4. Presione TRACE y use ◄ ► para ver los valores.

Valores atípicos

Cree una gráfica de caja modificada utilizando el procedimiento anterior u ordénelo de la manera siguiente:

1. Pulse STAT, seleccione **SortA** (orden ascendente) en el menú y pulse ENTER.
2. Introduzca el nombre de la lista que desea ordenar y pulse ENTER.
3. Para ver la lista ordenada, presione STAT, seleccione **Edit** y pulse ENTER.
4. Resalte la celda superior en una columna vacía, escriba el nombre de la lista y presione ENTER.
5. Utilice ▲ ▼ para examinar los valores mínimo y máximo y determinar si están muy alejados de los demás valores.

Excel

Resumen de 5 números

Use el procedimiento de **estadística descriptiva** dado al final de la sección 3-1 en la página 91.

Complemento XLSTAT

- Después del paso 3 en el procedimiento de estadística descriptiva, haga clic en la ficha **Outputs** y seleccione **Minimum**, **Maximum**, **1st Quartile**, **Median**, **3rd Quartile**. Haga clic en **OK**.

Excel

El complemento para análisis de datos proporciona sólo el mínimo, el máximo y la mediana. Para obtener los cuartiles utilice el siguiente procedimiento:

1. Haga clic en **Insert función** f_x, seleccione la categoría **Statistical** y seleccione la función **QUARTILE.INC**.
2. Introduzca el intervalo de valores de datos en el cuadro *Array*.
3. En el cuadro *Quart*, ingrese **0** para encontrar el mínimo, **1** para encontrar el primer cuartil y **2**, **3**, **4** para encontrar los valores restantes.

Gráficas de caja

Complemento XLSTAT (Requerido)

1. Haga clic en la ficha **XLSTAT** en la barra de opciones y, después haga clic en **Describing Data**.
2. Seleccione **Descriptive Statistics** en el menú desplegable.
3. Marque la casilla de **Quantitative Data** e introduzca el intervalo de datos deseado. Si selecciona dos o más columnas, se generarán varias gráficas. Si la primera fila de datos contiene una etiqueta, también marque la casilla **Sample labels**.
4. Haga clic en la ficha **Options** y confirme que la casilla **Charts** esté marcada.
5. Haga clic en la pestaña **Charts (1)** y marque la casilla **Box plots** en *Quantitative Data*.
6. Haga clic en **OK**.

Valores atípicos

Cree una gráfica de caja modificada utilizando el procedimiento de XLSTAT anterior u ordénela de la siguiente manera:

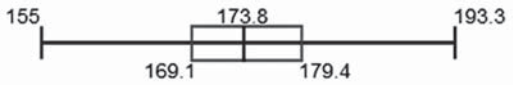
1. Haga clic en la ficha **Data** del menú superior y seleccione el intervalo de valores de datos deseado.
2. Haga clic en el botón **Sort Smallest to Largest** (A → Z) en la barra.
3. Examine los valores mínimo y máximo para determinar si están muy alejados de los demás valores.

3-3 Habilidades y conceptos básicos

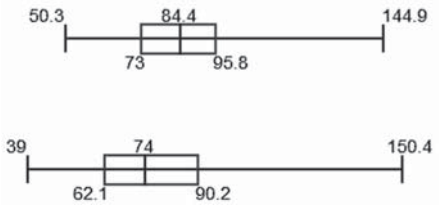
Conocimiento estadístico y pensamiento crítico

1. Puntuaciones z LeBron James, uno de los jugadores de baloncesto más exitosos de todos los tiempos, tiene una estatura de 6 pies y 8 pulgadas, o 203 cm. Con base en estadísticos del conjunto de datos 1 "Datos corporales" en el apéndice B, su estatura se convierte en la puntuación z de 4.07. ¿Cuántas desviaciones estándar está su estatura por arriba de la media?

2. Estaturas El diagrama de caja que se muestra a continuación es el resultado de las estaturas (en cm) de los varones listados en el conjunto de datos 1 "Datos corporales" del apéndice B. ¿Qué nos dicen los números en esa gráfica de caja?



3. Comparación de gráficas de caja Consulte las gráficas de caja que se muestran a continuación, las cuales están dibujadas en la misma escala. Una gráfica representa el peso de los hombres y la otra el peso de las mujeres. ¿Qué gráfica de caja representa el peso de las mujeres? Explique.



4. Puntuaciones z Si su calificación en su próximo examen de estadística se convierte en una puntuación z, ¿cuál de las siguientes puntuaciones prefiere: -2.00, -1.00, 0, 1.00, 2.00? ¿Por qué?

Puntuaciones z. ***En los ejercicios 5 a 8, exprese todas las puntuaciones z con dos decimales.***

5. Velocidades de datos en ATL Para las velocidades de datos en aeropuertos de Verizon (Mbps) listadas en el conjunto de datos 32 "Velocidades de datos en aeropuertos" del apéndice B, se midió la velocidad más alta de 77.8 Mbps en el aeropuerto internacional de Atlanta (ATL). La lista completa de 50 velocidades de datos para Verizon tiene una media de $\bar{x} = 17.60$ Mbps y una desviación estándar de $s = 16.02$ Mbps.

a. ¿Cuál es la diferencia entre la velocidad de datos para Verizon en el aeropuerto internacional de Atlanta y la media de todas las velocidades de datos para Verizon?

b. ¿Cuántas desviaciones estándar representa esto [la diferencia encontrada en el inciso (a)]?

c. Convierta la velocidad de datos para Verizon en el aeropuerto internacional de Atlanta a una puntuación z.

d. Si consideramos que las velocidades de datos que se convierten a puntuaciones z entre -2 y 2 no son significativamente bajas ni significativamente altas, ¿es la velocidad de Verizon en Atlanta significativa?

6. Velocidades de datos en PHL Repita el ejercicio anterior usando la velocidad de datos para Verizon, de 0.8 Mbps en el Aeropuerto Internacional de Filadelfia (PHL).

7. Pulsos de mujeres En el conjunto de datos 1 "Datos corporales" del apéndice B, se listan pulsos de mujeres. El pulso más bajo es de 36 latidos por minuto, la media de los pulsos listados es $\bar{x} = 74.0$ latidos por minuto y su desviación estándar es $s = 12.5$ latidos por minuto.

a. ¿Cuál es la diferencia entre el pulso de 36 latidos por minuto y el pulso medio de las mujeres?

b. ¿Cuántas desviaciones estándar representa esto [la diferencia encontrada en el inciso (a)]?

c. Convierta el pulso de 36 latidos por minuto en una puntuación z.

d. Si consideramos que los pulsos que se convierten en puntuaciones z entre -2 y 2 no son significativamente bajos ni significativamente altos, ¿es significativo el pulso de 36 latidos por minuto?

8. Desechos de plástico El conjunto de datos 31 "Peso de la basura" en el apéndice B contiene los pesos (en lb) del plástico desechado por los hogares. El peso más alto es 5.28 libras, la media de todos los pesos es $\bar{x} = 1.911$ libras y la desviación estándar de los pesos es $s = 1.065$ lb.

a. ¿Cuál es la diferencia entre el peso de 5.28 lb y la media de los pesos?

b. ¿Cuántas desviaciones estándar representa esto [la diferencia encontrada en el inciso (a)]?

c. Convierta el peso de 5.28 lb en una puntuación z.

d. Si consideramos que los pesos que se convierten en puntuaciones z entre -2 y 2 no son significativamente bajos ni significativamente altos, ¿es significativo el peso de 5.28 lb?

Valores significativos. ***En los ejercicios 9 a 12, considere un valor significativamente bajo si su puntuación z es menor o igual a −2 o significativamente alto si su puntuación z es mayor o igual a 2.***

9. ACT El examen ACT se utiliza con el fin de evaluar la preparación para la universidad. En un año reciente, la media del ACT fue de 21.1 y la desviación estándar fue de 5.1. Identifique las puntuaciones del ACT que son significativamente bajas o significativamente altas.

10. MCAT En un año reciente, las puntuaciones en el Medical College Admission Test (MCAT) tuvieron una media de 25.2 y una desviación estándar de 6.4. Identifique las puntuaciones del MCAT que son significativamente bajas o significativamente altas.

11. Monedas de 25 centavos El conjunto de datos 29 "Pesos de monedas" lista pesos (en gramos) de monedas de 25 centavos fabricados después de 1964. Estos pesos tienen una media de 5.63930 g y una desviación estándar de 0.06194 g. Identifique los pesos que son significativamente bajos o significativamente altos.

12. Diseño de asientos de avión En el proceso de diseño de asientos de avión se encontró que los hombres tienen un ancho de cadera con una media de 36.6 cm y una desviación estándar de 2.5 cm (con base en datos de la encuesta antropométrica de Gordon Clauser *et al.*). Identifique el ancho de cadera de los hombres que son significativamente bajos o significativamente altos.

Comparación de valores. ***En los ejercicios 13 a 16, use puntuaciones z para comparar los valores dados.***

13. Los hombres más altos y más bajos El hombre vivo más alto en el momento en que se escribió esto es Sultan Kosen, que tenía una estatura de 251 cm. El hombre vivo más bajo era Chandra Bahadur Dangi, con una estatura de 54.6 cm. Las estaturas de los hombres tienen una media de 174.12 cm y una desviación estándar de 7.10 cm. ¿Cuál de estos dos hombres tiene la estatura más extrema?

14. Conteo de glóbulos rojos Según el conjunto de datos 1 "Datos corporales" del apéndice B, los varones tienen conteos de glóbulos rojos con una media de 4.719 y una desviación estándar de 0.490, mientras que las mujeres tienen un conteo de glóbulos rojos con media de 4.349 y una desviación estándar de 0.402. ¿Quién tiene el conteo más alto en relación con la muestra de la cual provino: un hombre con un conteo de 5.58 o una mujer con un conteo de 5.23? Explique.

15. Pesos al nacer Según el conjunto de datos 4 "Nacimientos" del apéndice B, los varones recién nacidos tienen pesos con una media de 3272.8 g y una desviación estándar de 660.2 g. Las bebés recién nacidas tienen pesos con una media de 3037.1 g y una desviación estándar de 706.3 g. ¿Quién tiene el peso más extremo en relación con el grupo de donde provienen: un varón que pesó 1500 g o una bebé que pesó 1500 g?

16. Premios Oscar En la 87ª entrega de Premios de la Academia, Eddie Redmayne ganó el Oscar al mejor actor a la edad de 33 años y Julianne Moore el de mejor actriz a la edad de 54. Para todos los mejores actores, la edad media es 44.1 años y la desviación estándar es 8.9 años. Para todas las mejores actrices, la edad media es de 36.2 años y la desviación estándar es de 11.5 años. (Todas las edades se refieren al momento de la ceremonia de entrega de premios). En relación con sus géneros, ¿quién tenía la edad más extrema al ganar el Oscar: Eddie Redmayne o Julianne Moore? Explique.

Percentiles. ***En los ejercicios 17 a 20, utilice las siguientes velocidades de datos en aeropuertos para teléfonos celulares (Mbps) de Sprint. Encuentre el percentil correspondiente a la velocidad de datos dada.***

0.2	0.3	0.3	0.3	0.3	0.3	0.3	0.4	0.4	0.4
0.5	0.5	0.5	0.5	0.5	0.6	0.6	0.7	0.8	1.0
1.1	1.1	1.2	1.2	1.6	1.6	2.1	2.1	2.3	2.4
2.5	2.7	2.7	2.7	3.2	3.4	3.6	3.8	4.0	4.0
5.0	5.6	8.2	9.6	10.6	13.0	14.1	15.1	15.2	30.4

17. 2.4 Mbps **18.** 13.0 Mbps **19.** 0.7 Mbps **20.** 9.6 Mbps

En los ejercicios 21 a 28, use la misma lista de velocidades de datos en aeropuertos para Sprint (Mbps) dada para los ejercicios 17 a 20. Encuentre el percentil o cuartil indicado.

21. P_{60} **22.** Q_1 **23.** Q_3 **24.** P_{40} **25.** P_{50} **26.** P_{75} **27.** P_{25} **28.** P_{85}

Gráficas de caja. ***En los ejercicios 29 a 32, utilice los datos dados para construir una gráfica de caja e identifique el resumen de 5 números.***

29. Citas rápidas Las siguientes son las calificaciones que han dado los varones a las mujeres en un experimento que implica citas rápidas.

2.0 3.0 4.0 5.0 6.0 6.0 7.0 7.0 7.0 7.0 7.0 7.0 8.0 8.0 8.0 8.0 9.0 9.5 10.0 10.

30. Radiación del teléfono celular A continuación se listan las tasas de absorción de radiación medidas (en W/kg) correspondientes a los siguientes teléfonos celulares: iPhone 5S, BlackBerry Z30, Sanyo Vero, Optimus V, Droid Razr, Nokia N97, Samsung Vibrant, Sony Z750a, Kyocera Kona, LG G2 y Virgin Mobile Supreme. Los datos son de la Comisión Federal de Comunicaciones.

1.18 1.41 1.49 1.04 1.45 0.74 0.89 1.42 1.45 0.51 1.38

31. Radiación en dientes de bebés A continuación se listan las cantidades de estroncio-90 (en milibecquerelios o mBq) en una muestra aleatoria simple de dientes de leche obtenida de residentes de Pensilvania nacidos después de 1979 (con base en datos de "An Unexpected Rise in Strontium-90 in U.S. Deciduous Teeth in the 1990s", de Mangano *et al*., *Science of the Total Environment*).

128 130 133 137 138 142 142 144 147 149 151 151 151 155
156 161 163 163 166 172

32. Medidas de presión arterial Catorce estudiantes diferentes de segundo año de medicina en el Hospital Bellevue midieron la presión arterial de la misma persona. A continuación se indican las lecturas sistólicas (mm Hg).

138 130 135 140 120 125 120 130 130 144 143 140 130 150

Gráficas de cajas de los conjuntos de datos grandes en el apéndice B. ***En los ejercicios 33 a 36, utilice los conjuntos de datos dados en el apéndice B. Utilice las gráficas de caja para comparar los dos conjuntos de datos.***

33. Pulsos Utilice la misma escala para construir gráficas de caja para los pulsos de hombres y mujeres del conjunto de datos 1 "Datos corporales" en el apéndice B.

34. Edades de los ganadores del Oscar Use la misma escala para construir gráficas de cajas para las edades de las mejores actrices y mejores actores del conjunto de datos 14 "Edades de ganadores del Oscar" en el apéndice B.

35. IMC Utilice los índices de masa corporal (IMC) para hombres y mujeres que se listan en el conjunto de datos 1 "Datos corporales".

36. Plomo e IQ Utilice la misma escala para construir las gráficas de caja para las puntuaciones de IQ completas (IQF) para el grupo de nivel bajo de plomo (grupo 1) y el grupo de nivel alto de plomo (grupo 3) en el conjunto de datos 1 "IQ y plomo" en el apéndice B.

3-3 Más allá de lo básico

37. Valores atípicos y gráficas de caja modificadas Repita el ejercicio 33 "Pulsos" utilizando gráficas de caja modificadas. Identifique los valores atípicos definidos en la parte 2 de esta sección.

Examen rápido del capítulo

1. Media de sueño Como parte de la Encuesta Nacional de Exámenes de Salud y Nutrición, se preguntó a los sujetos cuánto tiempo durmieron la noche anterior, y se reportaron los siguientes tiempos (horas): 8, 7, 5, 7, 4, 7, 6, 7, 8, 8, 8, 6. Encuentre la media.

2. Mediana de sueño ¿Cuál es la mediana de los valores muestrales listados en el ejercicio 1?

3. Moda de sueño ¿Cuál es la moda de los valores muestrales listados en el ejercicio 1?

4. Varianza de sueño La desviación estándar de los valores muestrales en el ejercicio 1 es de 1.3 horas. ¿Cuál es la varianza (incluyendo las unidades)?

5. Valores atípicos de sueño Si un tiempo de sueño de 0 horas se incluye en los datos muestrales dados en el ejercicio 1, ¿es un valor atípico? ¿Por qué sí o por qué no?

6. Puntuación z de sueño Una muestra mayor de 50 tiempos de sueño (horas) tiene una media de 6.3 horas y una desviación estándar de 1.4 horas. ¿Cuál es la puntuación z para un tiempo de sueño de 5 horas?

7. Q_3 de sueño Para una muestra de 80 tiempos de sueño, ¿aproximadamente cuántos de esos tiempos son menores que Q_3?

8. Resumen de 5 números de sueño Para una muestra de 100 tiempos de sueño, dé los nombres de los valores que constituyen el resumen de 5 números. (Los valores reales no se pueden identificar).

9. Estimación de s Una gran muestra de tiempos de sueño incluye valores que van desde un mínimo de 4 horas hasta un máximo de 10 horas. Utilice la regla práctica del rango para calcular la desviación estándar.

10. Notación de sueño Considere una muestra de los tiempos de sueño tomadas de la población adulta que vive en Alaska. Identifique los símbolos utilizados para la media de la muestra, la media de la población, la desviación estándar de la muestra, la desviación estándar de la población, la varianza de la muestra y la varianza de la población.

Ejercicios de repaso

1. Géiser Old Faithful A continuación se listan los errores de predicción (minutos) que son las diferencias entre los tiempos reales de erupción y los tiempos de erupción previstos. Los números positivos corresponden a erupciones que ocurrieron más tarde de lo previsto, y los números negativos corresponden a erupciones que ocurrieron antes del pronóstico. (Los datos provienen del conjunto de datos 23 "Old Faithful" en el apéndice B). Encuentre (a) la media; (b) la mediana; (c) la modo; (d) la mitad del rango; (e) el rango; (f) la desviación estándar; (g) la varianza; (h) Q_1; (i) Q_3

$$4 \quad -7 \quad 0 \quad 1 \quad -1 \quad 1 \quad -4 \quad -7 \quad 22 \quad 7 \quad -5 \quad 1$$

2. Puntuación z Usando los datos muestrales del ejercicio 1, encuentre la puntuación z correspondiente al error de predicción de 0 min. ¿Es ese error de predicción significativamente bajo o alto? ¿Por qué sí o por qué no?

3. Gráfica de caja Utilizando los mismos errores de predicción que se listan en el ejercicio 1, construya un diagrama de caja e incluya los valores del resumen de 5 números.

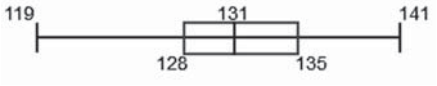
4. Códigos de emergencia En un análisis de las actividades que resultaron en lesiones cerebrales presentadas en las salas de emergencia de hospital, se identificaron las siguientes actividades mediante los códigos que se muestran entre paréntesis: ciclismo (12); fútbol (14); parque infantil (22); baloncesto (27); natación (40). Encuentre la media de 12, 14, 22, 27 y 40. ¿Qué tiene de erróneo este resultado?

5. Comparación de pesos al nacer El peso al nacer de una muestra de varones tiene una media de 3272.8 g y una desviación estándar de 660.2 g. El peso al nacer de una muestra de niñas tiene una media de 3037.1 g y una desviación estándar de 706.3 g (con base en el conjunto de datos 4 "Nacimientos" en el apéndice B). Cuando se considera entre los miembros del mismo sexo, ¿cuál bebé tiene el peso al nacer relativamente mayor: un varón con un peso al nacer de 3400 g o una niña con un peso al nacer de 3200 g? ¿Por qué?

6. Efectos de un valor atípico A continuación se listan los conteos de plaquetas (1000 células/μl) de los sujetos incluidos en el conjunto de datos 1 "Datos corporales". Identifique el valor atípico y luego comente el efecto que tiene sobre la media y la desviación estándar, mediante la determinación de los valores de esos estadísticos incluyendo el valor atípico y sin incluirlo.

263 206 185 246 188 191 308 262 198 253 646

7. Interpretación de una gráfica de caja A continuación se muestra una gráfica de caja de una muestra de 30 anchos máximos de cráneo (mm) medidos en cráneos egipcios de alrededor del año 4000 a.C. ¿Qué representan los números en la gráfica de caja?



8. Estimación de la desviación estándar A continuación se lista una muestra de tiempos de duración (segundos) de erupciones del géiser Old Faithful. Utilice la regla práctica del rango para estimar el valor de la desviación estándar de todos los tiempos de duración y compare el resultado con la desviación estándar de 33.7 segundos obtenida de una muestra de 2634 tiempos de duración.

226 228 247 247 253 256 250 254 229 242 250 241 226 240 117

Ejercicios de repaso acumulado

1. Arsénico en el arroz A continuación se listan las cantidades medidas (μg por porción) de arsénico en una muestra de porciones de arroz integral [datos de la Food and Drug Administration (FDA)]. Construya una distribución de frecuencias. Utilice una anchura de clase de 2 μg y use 0 μg como límite inferior de la primera clase.

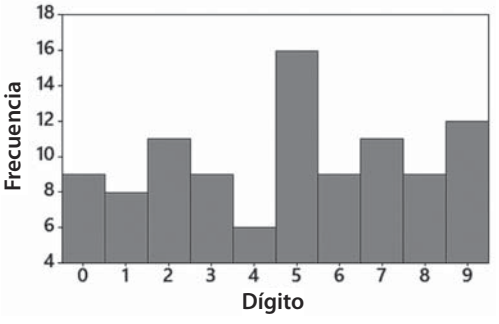
6.1 5.4 6.9 4.9 6.6 6.3 6.7 8.2 7.8 1.5 5.4 7.3

2. Histograma Utilice la distribución de frecuencias del ejercicio 1 para construir un histograma. Utilice los valores medios de clase para la escala horizontal.

3. Diagrama de tallo y hojas Utilice las cantidades de arsénico del ejercicio 1 para construir un diagrama de tallo y hojas.

4. Estadísticos descriptivos Utilice cantidades de arsénico en el ejercicio 1 y encuentre lo siguiente: (a) media, (b) mediana, (c) desviación estándar, (d) varianza y (e) rango. Incluya las unidades de medida apropiadas.

5. Histograma El histograma adjunto muestra los resultados de los dígitos de la lotería Florida Play 4. ¿Cuál es el error más importante en este histograma?



6. Distribución normal Examine la distribución mostrada en el histograma del ejercicio 5. ¿Parece que los datos muestrales provienen de una población con una distribución normal? ¿Por qué sí o por qué no?

Proyecto de tecnología

Palabras pronunciadas por hombres y mujeres Considere el conjunto de datos 24 "Conteos de palabras" en el apéndice B, que incluye los conteos de palabras pronunciadas por hombres y mujeres. Ese conjunto de datos incluye 12 columnas de datos, pero primero apile todos los conteos de palabras masculinas en una columna y todos los conteos de palabras femeninas en otra columna. A continuación, proceda a generar histogramas, o cualquier otra gráfica adecuada, y encuentre los estadísticos que le permiten comparar los dos conjuntos de datos. ¿Existen valores atípicos? ¿Ambos conjuntos de datos tienen propiedades que son básicamente iguales? ¿Hay diferencias significativas? ¿Cuál sería una consecuencia de tener diferencias significativas? Escriba un breve informe incluyendo sus conclusiones y gráficas de apoyo.

DE LOS DATOS A LA DECISIÓN

Humo de segunda mano

El conjunto de datos 12 "Humo pasivo y activo" en el apéndice B contiene medidas de cotinina en tres grupos de sujetos: (1) fumadores, (2) no fumadores expuestos al humo ambiental del tabaco y (3) no fumadores no expuestos al humo ambiental del tabaco. La cotinina es un indicador de la absorción de nicotina.

Pensamiento crítico

Utilice los métodos de este capítulo para explorar y comparar las medidas de cotinina en los tres grupos. ¿Hay diferencias notables? ¿Existen valores atípicos? ¿Qué concluye usted acerca de los efectos que los fumadores tienen sobre los no fumadores? Escriba un informe breve de sus conclusiones, y proporcione evidencia estadística de apoyo.

Actividades en equipo

1. Actividad en clase En clase, cada estudiante debe registrar dos pulsos contando el número de latidos en 1 minuto. El primer pulso se debe medir mientras el estudiante está sentado, y el segundo pulso debe medirse mientras el estudiante está de pie. Utilice los métodos de este capítulo para comparar los resultados. ¿Los hombres y las mujeres parecen tener diversos pulsos? ¿Los pulsos medidos mientras están sentados parecen ser diferentes de los pulsos medidos mientras están de pie?

2. Actividades fuera de clase En el artículo "Weighing Anchors" en la revista *Omni*, el autor John Rubin observó que cuando la gente estima un valor, su estimación suele estar "anclada" (o influenciada por) un número precedente, incluso si ese número anterior no está relacionado con la cantidad que se estima. Para demostrar esto, pidió a las personas que dieran una estimación rápida del valor de $8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$. La media de las respuestas dadas fue 2250, pero cuando el orden de los números fue invertido, la media se convirtió en 512. Rubin explicó que cuando comenzamos los cálculos con números más grandes (como en $8 \times 7 \times 6$), nuestras estimaciones tienden a ser más grandes. Señaló que tanto 2250 como 512 están muy por debajo del producto correcto, 40,320. El artículo sugiere que hay cifras irrelevantes que pueden influir en las evaluaciones de bienes raíces, las estimaciones de los valores de los automóviles y las estimaciones de la probabilidad de una guerra nuclear.

Realice un experimento para probar esta teoría. Seleccione algunos sujetos y pídales que estimen rápidamente el valor de

$$8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$$

Luego seleccione otros sujetos y pídales que estimen rápidamente el valor de

$$1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8$$

Registre las estimaciones junto con el orden particular utilizado. Diseñe cuidadosamente el experimento para que las condiciones sean uniformes y los dos grupos muestrales se seleccionen de manera que se minimice cualquier sesgo. No describa la teoría a los sujetos hasta después de haber proporcionado sus estimaciones. Compare los dos conjuntos de resultados de la muestra utilizando los métodos de este

capítulo. Proporcione un informe impreso que incluya los datos recopilados, los métodos detallados que se usaron, el método de análisis, las gráficas y/o estadísticos pertinentes y una declaración de conclusiones. Incluya una crítica del experimento, con razones por las cuales los resultados pudieron no ser correctos, y describa maneras en las cuales el experimento podría ser mejorado.

3. Actividad en clase Complete la siguiente tabla y compare la variación en las velocidades de datos para teléfono inteligente. Consulte el conjunto de datos 32 "Velocidades de datos en aeropuertos".

	Verizon	Sprint	AT&T	T-Mobile
Desviación estándar				
Rango				

4. Actividad fuera de clase Registre los tiempos que los automóviles están estacionados en una bomba de gasolina, y describa las características importantes de esos tiempos.

5. Actividad fuera de clase Varios sitios web, como www.gasbuddy.com, están diseñados para proporcionar una lista de los precios locales de la gasolina. Obtenga una lista de tales precios y explore los datos usando los métodos de este capítulo y el capítulo 2.

6. Actividad fuera de clase El conjunto de datos 28 "Galletas con chispas de chocolate" en el apéndice B incluye conteos de chispas de chocolate en cinco diferentes marcas de galletas. Obtenga su propia muestra de galletas de chispas de chocolate y proceda a contar el número de chispas en cada galleta. Utilice los datos para generar un histograma y cualquier otra gráfica adecuada. Encuentre los estadísticos descriptivos. Compare sus conteos de chispas de chocolate con los dados en el conjunto de datos 28. ¿Hay alguna diferencia? Explique.

7. Actividad fuera de clase El apéndice B incluye muchos conjuntos de datos reales e interesantes. En grupos de tres o cuatro estudiantes, seleccionen un conjunto de datos del apéndice B y analícenlos usando los métodos discutidos hasta ahora en este libro. Escriban un breve informe resumiendo las conclusiones clave.

8. Actividad fuera de clase En grupos de tres o cuatro estudiantes, recolecten un conjunto de valores de datos originales en el intervalo o nivel de relación de medida. Proporcionen lo siguiente: (1) una lista de los valores muestrales; (2) los resultados de un software para estadísticos descriptivos y gráficas; y (3) una descripción escrita de la naturaleza de los datos, el método de recolección y las características importantes.

5.4.3. Técnicas estadísticas para el análisis de datos

Para el desarrollo del subtema **Técnicas estadísticas para el análisis de datos**, se consideran aquellos procedimientos estadísticos que permiten **organizar, describir, analizar e interpretar los datos obtenidos en una investigación cuantitativa**, incluyendo la estadística descriptiva y la estadística inferencial. Se abordan las distribuciones de frecuencias, las medidas de tendencia central y de variabilidad, así como las pruebas estadísticas para la comprobación de hipótesis, entre ellas las pruebas paramétricas y no paramétricas. Estos contenidos permiten seleccionar e interpretar técnicas estadísticas de acuerdo con los objetivos y características de la investigación.

N.°	Libros	Capítulo	Página	Contenido
1	*Metodología de la Investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**p. 270**	**Análisis de datos cuantitativos:** procedimiento general para el análisis de los datos obtenidos en una investigación cuantitativa.
2	*Metodología de la Investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**pp. 282–285**	**Estadística descriptiva:** distribución de frecuencias y diferentes formas de presentación de los datos.
3	*Metodología de la Investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**pp. 286–290**	**Medidas de tendencia central y variabilidad:** media, mediana, moda, medidas de variabilidad, varianza e interpretación de las medidas estadísticas.
4	*Metodología de la Investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**pp. 293–294**	**Puntuaciones z, razones y tasas:** procedimientos estadísticos complementarios para el análisis descriptivo de los datos.
5	*Metodología de la Investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**pp. 299–304**	**Estadística inferencial y prueba de hipótesis:** inferencia de la muestra a la población, distribución muestral, nivel de significancia y prueba de hipótesis.
6	*Metodología de la Investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**pp. 304–314**	**Pruebas estadísticas paramétricas:** coeficiente de correlación de Pearson, regresión lineal, prueba *t*, tamaño del efecto, diferencia de proporciones y ANOVA de un factor.
7	*Metodología de la Investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**pp. 316–318**	**Estadística multivariada y análisis no paramétricos:** introducción a procedimientos multivariados y pruebas estadísticas no paramétricas.

8	*Metodología de la Investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**pp. 318–324**	**Pruebas no paramétricas:** Chi cuadrada, coeficientes de correlación e independencia para tabulaciones cruzadas y otros coeficientes de correlación.
9	*Metodología de la Investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**pp. 322–324**	**Correlación por rangos:** coeficientes de correlación de Spearman y Kendall y otros coeficientes.
10	*Metodología de la Investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**pp. 324–327**	**Selección y aplicación de procedimientos estadísticos:** vista general de los procedimientos o pruebas estadísticas, análisis adicionales y preparación de los resultados para su presentación.

5.4.3. Técnicas estadísticas para el análisis de datos

Para el desarrollo del subtema **Técnicas estadísticas para el análisis de datos**, se consideran aquellos procedimientos estadísticos que permiten analizar, describir, comparar y establecer relaciones entre los datos obtenidos en una investigación. Se incluyen técnicas de estadística descriptiva, estimación de parámetros, pruebas de hipótesis, comparación de muestras, correlación y regresión, así como procedimientos de análisis de varianza y pruebas no paramétricas. Estos contenidos permiten seleccionar y aplicar técnicas estadísticas de acuerdo con las características de los datos y los objetivos del análisis.

N.°	Libros	Capítulo	Página	Contenido
1	*Estadística* — Mario F. Triola	**Capítulo 3. Descripción, exploración y comparación de datos**	**p. 80**	Medidas de tendencia central, medidas de variación y medidas de posición relativa como procedimientos para describir y analizar cuantitativamente los datos.
2	*Estadística* — Mario F. Triola	**Capítulo 7. Estimación de parámetros y determinación del tamaño de la muestra**	**p. 297**	Estimación de una proporción poblacional, de un promedio poblacional y de una desviación estándar o varianza poblacional.
3	*Estadística* — Mario F. Triola	**Capítulo 8. Pruebas de hipótesis**	**p. 356**	Fundamentos de las pruebas de hipótesis y procedimientos para contrastar hipótesis sobre proporciones, medias y varianzas.
4	*Estadística* — Mario F. Triola	**Capítulo 9. Inferencias a partir de dos muestras**	**p. 414**	Técnicas de inferencia estadística para comparar dos proporciones, dos medias de muestras independientes o dependientes y dos varianzas.
5	*Estadística* — Mario F. Triola	**Capítulo 10. Correlación y regresión**	**p. 468**	Análisis de correlación, regresión, intervalos de predicción y regresión múltiple para estudiar relaciones entre variables.
6	*Estadística* — Mario F. Triola	**Capítulo 11. Bondad de ajuste y tablas de contingencia**	**p. 533**	Procedimientos de bondad de ajuste y análisis de tablas de contingencia para estudiar la asociación y distribución de variables categóricas.
7	*Estadística* — Mario F. Triola	**Capítulo 12. Análisis de varianza**	**p. 566**	ANOVA de un factor y ANOVA de dos factores para analizar y comparar diferencias entre medias de grupos.
8	*Estadística* — Mario F. Triola	**Capítulo 13. Pruebas no paramétricas**	**p. 597**	Técnicas no paramétricas para el análisis estadístico, incluyendo prueba del signo, pruebas de Wilcoxon, Kruskal-Wallis y correlación de rangos.

5.4.3. Técnicas estadísticas para el análisis de datos

Para el desarrollo del subtema **Técnicas estadísticas para el análisis de datos**, se consideran aquellos procedimientos estadísticos que permiten analizar los datos obtenidos en una investigación mediante medidas estadísticas, estimación, pruebas de hipótesis, comparación de grupos, análisis de asociación y análisis de relaciones entre variables.

N.°	Libros	Capítulo	Página	Contenido
1	*Estadística para Administración y Economía*	Capítulo 3. Medidas de tendencia central y dispersión	**p. 57**	Medidas de tendencia central y dispersión para el análisis estadístico de los datos.
2	*Estadística para Administración y Economía*	Capítulo 7. Estimación	**p. 273**	Estimación puntual e intervalos de confianza para realizar inferencias estadísticas.
3	*Estadística para Administración y Economía*	Capítulo 8. Prueba de hipótesis: prueba de una sola muestra	**p. 319**	Pruebas de hipótesis para medias y proporciones.
4	*Estadística para Administración y Economía*	Capítulo 9. Prueba de hipótesis: pruebas de dos muestras	**p. 359**	Comparación estadística de medias y proporciones entre dos muestras.
5	*Estadística para Administración y Economía*	Capítulo 11. Ji-cuadrada y análisis de varianza	**p. 447**	Prueba de independencia, bondad de ajuste y análisis de varianza.
6	*Estadística para Administración y Economía*	Capítulo 12. Regresión simple y correlación	**p. 509**	Análisis de correlación y regresión para estudiar relaciones entre variables.
7	*Estadística para Administración y Economía*	Capítulo 14. Métodos no paramétricos	**p. 621**	Técnicas no paramétricas para el análisis estadístico de datos.

capítulo 3

MEDIDAS DE TENDENCIA CENTRAL Y DISPERSIÓN EN DISTRIBUCIONES DE FRECUENCIAS

Objetivos

- Utilizar la estadística sumaria para describir una colección de datos
- Utilizar la media, la mediana y la moda para describir cómo se "aglutinan" los datos
- Utilizar el rango, la varianza y la desviación estándar para describir cómo se "dispersan" los datos
- Examinar los análisis de datos exploratorios, basados en el uso de la computadora, para conocer otras formas útiles de resumir los datos

Contenido del capítulo

3.1 Estadística sumaria 58
3.2 Una medida de tendencia central: la media aritmética 60
3.3 Una segunda medida de tendencia central: la media ponderada 69
3.4 Una tercera medida de tendencia central: la media geométrica 74
3.5 Una cuarta medida de tendencia central: la mediana 77
3.6 Una medida final de tendencia central: la moda 84
3.7 Dispersión: por qué es importante 89
3.8 Rangos: medidas de dispersión útiles 91
3.9 Dispersión: medidas de desviación promedio 96
3.10 Dispersión relativa: el coeficiente de variación 107
3.11 Análisis exploratorio de datos (AED) 112
- Estadística en el trabajo 116
- Ejercicio de base de datos computacional 117
- Términos introducidos en el capítulo 3 118
- Ecuaciones introducidas en el capítulo 3 119
- Ejercicios de repaso 121

El vicepresidente de mercadotecnia de una cadena de restaurantes de comida rápida está estudiando el desarrollo de las ventas de las 100 sucursales que se encuentran en el distrito oriental y ha elaborado la siguiente distribución de frecuencias para las ventas anuales:

Ventas (miles)	Frecuencia	Ventas (miles)	Frecuencia
700- 799	4	1,300-1,399	13
800- 899	7	1,400-1,499	10
900- 999	8	1,500-1,599	9
1,000-1,099	10	1,600-1,699	7
1,100-1,199	12	1,700-1,799	2
1,200-1,299	17	1,800-1,899	1

El vicepresidente desea comparar las ventas del distrito oriental con las ventas de otros tres distritos del país. Para llevar a cabo esto, hará un resumen de la distribución, poniendo especial cuidado en el acopio de información sobre la tendencia central de los datos. En este capítulo analizaremos también cómo se puede medir la variabilidad de una distribución y, por tanto, cómo obtener una percepción mucho mejor de los datos. ■

3.1 Estadística sumaria

Estadística sumaria, tendencia central y dispersión

En el capítulo 2 construimos tablas y gráficas a partir de una colección de datos sin procesar. Los "retratos" resultantes de las distribuciones de frecuencias ilustraron tendencias y patrones de los datos. En casi todos los casos, sin embargo, teníamos necesidad de medidas más exactas. En estos casos, podemos usar los números que constituyen la *estadística sumaria* para describir las características del conjunto de datos.

Dos de estas características son de particular importancia para los responsables de tomar decisiones: la *tendencia central* y la *dispersión*.

Punto medio de un conjunto de datos

Tendencia central La tendencia central se refiere al punto medio de una distribución. Las medidas de tendencia central se conocen también como *medidas de posición*. En la figura 3-1, la posición central de la curva B está a la derecha de las posiciones centrales de las curvas A y C. Observe que la posición central de la curva A es la misma que la de la curva C.

Separación de un conjunto de datos

Dispersión La *dispersión* se refiere a la separación de los datos en una distribución, es decir, al grado en que las observaciones se separan. Note que la curva A de la figura 3-2 tiene una mayor separación o dispersión que la curva B.

Existen otras dos características de los conjuntos de datos que proporcionan información útil: el *sesgo* y la *curtosis*. Aunque la derivación de la estadística específica para medir dichas caracterísi-

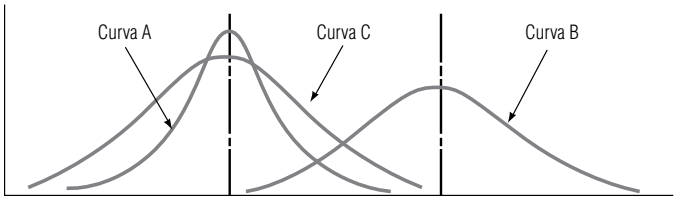


FIGURA 3-1

Comparación de la posición central de tres curvas

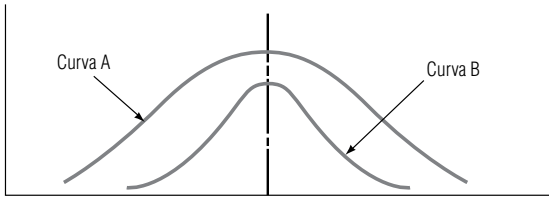


FIGURA 3-2

Comparación de la dispersión de dos curvas

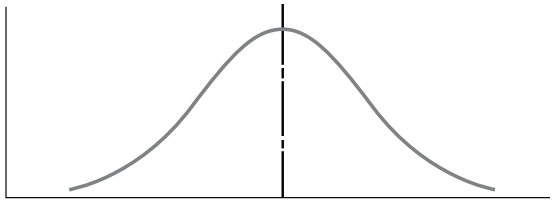


FIGURA 3-3

Curva simétrica

cas está más allá de los objetivos de este texto, nos será útil tener un conocimiento general de su significado.

Simetría de un conjunto de datos

Sesgo Las curvas que representan los datos puntuales de un conjunto de datos pueden ser simétricas o sesgadas. Las curvas *simétricas*, como la de la figura 3-3, tienen una forma tal que una línea vertical que pase por el punto más alto de la curva dividirá su área en dos partes iguales. Cada parte es una imagen de espejo de la otra.

Sesgo de un conjunto de datos

Las curvas A y B de la figura 3-4 son curvas *sesgadas*. Están sesgadas porque los valores de su distribución de frecuencias se concentran en el extremo inferior o en el superior de la escala de medición del eje horizontal. Estos valores no están igualmente distribuidos. La curva A está sesgada a la derecha (o *positivamente* sesgada), debido a que va disminuyendo poco a poco hacia el extremo derecho de la escala. La curva B es exactamente opuesta. Está sesgada a la izquierda (*negativamente* sesgada), ya que disminuye poco a poco si la recorremos hacia el extremo inferior de la escala.

La curva A podría representar la distribución de frecuencias del número de días que un producto se encuentra en existencia en un negocio de venta de fruta al mayoreo. La curva estaría sesgada a la derecha, con muchos valores en el extremo izquierdo y pocos en el extremo derecho, debido a que el inventario debe agotarse rápidamente. De manera análoga, la curva B podría representar la frecuencia del número de días que requiere un agente de bienes raíces para vender una casa. Estaría sesgada hacia la izquierda, con muchos valores en el extremo derecho de la escala y pocos en el izquierdo, debido a que el inventario de casas se coloca muy lentamente.

Agudeza de un conjunto de datos

Curtosis Cuando medimos la *curtosis* de una distribución, estamos midiendo qué tan puntiaguda es. En la figura 3-5, por ejemplo, las curvas A y B difieren entre sí sólo en que una tiene un pico más pronunciado que la otra. Tienen la misma posición central y la misma dispersión, y ambas son simétricas. Los estadísticos dicen que tienen un grado diferente de curtosis.

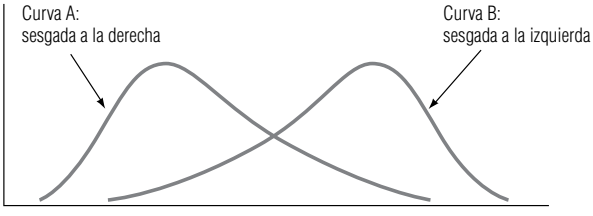


FIGURA 3-4

Comparación de dos curvas sesgadas

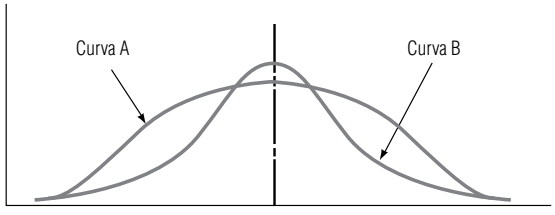


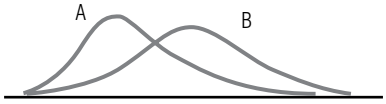
FIGURA 3-5

Dos curvas con la misma posición central pero diferente curtosis

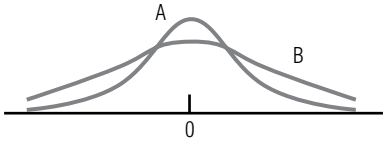
Ejercicios 3.1

Conceptos básicos

- **3-1** Trace tres curvas, todas simétricas, pero con diferente dispersión.
- **3-2** Trace tres curvas, todas simétricas y con la misma dispersión, pero con las siguientes posiciones centrales:
  a) 0.0 b) 1.0 c) −1.0
- **3-3** Trace una curva que pudiera ser una buena representación de las calificaciones en un examen de estadística de un grupo mal preparado, y también la de un grupo bien preparado.
- **3-4** Para las distribuciones siguientes, indique cuál de ellas
  a) tiene el valor promedio más grande.
  b) es más probable que produzca un valor pequeño que uno grande.
  c) es la mejor representación de la distribución de edades de los asistentes a un concierto de rock.
  d) es la mejor representación de la distribución de los tiempos de espera de pacientes en el consultorio de un médico.



Para las siguientes dos distribuciones, indique cuál de ellas, si alguna,
  e) tiene valores distribuidos más uniformemente a través del intervalo de valores posibles.
  f) es más probable que produzca un valor cercano a cero.
  g) tiene una probabilidad más alta de producir valores positivos que negativos.



- **3-5** Si las dos curvas siguientes representan la distribución de los resultados de un grupo de estudiantes en dos exámenes, ¿cuál examen parece haber sido más difícil para los estudiantes?

3.2 Una medida de tendencia central: la media aritmética

Casi siempre, cuando nos referimos al "promedio" de algo, estamos hablando de la media aritmética. Esto es cierto en casos como la temperatura invernal promedio en la ciudad de Nueva York, la vida promedio de la batería del *flash* de una cámara o la producción promedio de maíz en una hectárea de tierra.

La media aritmética

La tabla 3-1 presenta datos que describen el número de días que los generadores de una planta de energía de Lake Ico se encuentran fuera de servicio debido a mantenimiento normal o por alguna falla. Para encontrar la media aritmética, sumamos los valores y dividimos el resultado entre el número de observaciones:

$$\text{Media aritmética} = \frac{7 + 23 + 4 + 8 + 2 + 12 + 6 + 13 + 9 + 4}{10}$$

$$= \frac{88}{10}$$

$$= 8.8 \text{ días}$$

Tabla 3-1
Tiempo sin funcionar de los generadores de la estación de Lake Ico

Generador	1	2	3	4	5	6	7	8	9	10
Días fuera de servicio	7	23	4	8	2	12	6	13	9	4

En el periodo de un año, los generadores estuvieron fuera de servicio un promedio de 8.8 días. Con esta cifra, el administrador de la planta de energía tiene una medida sencilla y razonable del comportamiento de *todos* sus generadores.

Símbolos convencionales

Las características de una muestra se conocen como estadísticos

Las características de una población se llaman parámetros

Para escribir ecuaciones de este tipo de medidas de las distribuciones de frecuencias, necesitamos aprender la notación matemática que utilizan los especialistas en estadística. Una *muestra* de una población consiste en n observaciones (con n minúscula) con una media de $\bar{x}$ (x barra). Recuerde que las medidas calculadas para una muestra se conocen como *estadísticos*.

La notación es diferente cuando calculamos medidas para la *población* entera, es decir, para el grupo que contiene a todos los elementos que estamos describiendo. La media de una población se simboliza con μ, que es la letra griega *mu*. El número de elementos de una población se denota con la letra mayúscula cursiva N. Por lo general, en estadística se usan letras del alfabeto latino para simbolizar la información de las muestras y letras griegas para referirnos a la información de las poblaciones.

Cálculo de la media a partir de datos no agrupados

Encontrar las medias de la población y de la muestra

En el ejemplo, el promedio de 8.8 días sería μ (la media de la población) si la población de generadores fuera exactamente 10. Sería $\bar{x}$ (la media de la muestra), si los 10 generadores fueran una muestra tomada de una población mayor de ellos. Para escribir las fórmulas correspondientes a estas dos medias, combinamos los símbolos matemáticos y los pasos que utilizamos para determinar la media aritmética. Si se suman los valores de las observaciones y esta suma se divide entre el número de observaciones, obtendremos:

Media aritmética de la población

$$\mu = \frac{\Sigma x}{N} \qquad [3\text{-}1]$$

Σx: Suma de los valores de todas las observaciones

N: Número de elementos de la población

y

Media aritmética de la muestra

$$\bar{x} = \frac{\Sigma x}{n} \qquad [3\text{-}2]$$

Σx: Suma de los valores de todas las observaciones

n: Número de elementos de la muestra

Debido a que μ es la *media aritmética de la población*, usamos N para indicar que se divide entre el número de observaciones o elementos de la población. Del mismo modo, $\bar{x}$ es la media aritmética de

Tabla 3-2
Resultados del examen de aptitud académica

Estudiante	1	2	3	4	5	6	7
Aumento	9	7	7	6	4	4	2

la muestra, y n es el número de observaciones de la muestra. La letra griega sigma, Σ, indica que todos los valores de x se suman.

Otro ejemplo: en la tabla 3-2 se presenta la lista del aumento en puntos porcentuales en los resultados de siete estudiantes que tomaron un curso de preparación para el examen oral de aptitud escolar.

Calculamos la media de esta muestra de siete estudiantes de la manera siguiente:

$$\bar{x} = \frac{\Sigma x}{n} \qquad [3\text{-}2]$$

$$= \frac{9 + 7 + 7 + 6 + 4 + 4 + 2}{7}$$

$$= \frac{39}{7}$$

$$= 5.6 \text{ puntos por estudiante} \longleftarrow \text{Media de la muestra}$$

Manejo de datos no agrupados

Observe que para calcular esta media, sumamos todas las observaciones. Los especialistas en estadística se refieren a este tipo de datos como datos *no agrupados*. Los cálculos no fueron difíciles, pues nuestro tamaño de muestra era pequeño. Pero suponga que debe trabajar con el peso de 5,000 cabezas de ganado y prefiere no sumar por separado cada uno de los datos; o suponga que tiene acceso sólo a la distribución de frecuencias de los datos y no a cada observación individual. En estos casos, se requiere una manera distinta de calcular la media aritmética.

Cálculo de la media a partir de datos agrupados

Manejo de datos agrupados

Una distribución de frecuencias consta de datos agrupados en clases. Cada valor de una observación cae dentro de alguna de las clases. A diferencia del ejemplo del examen de aptitud, no conocemos el valor individual de cada observación. Suponga que tenemos una distribución de frecuencias (ilustrada en la tabla 3-3) del saldo promedio mensual de la cuenta de cheques de 600 clientes de una sucursal bancaria. A partir de la información de la tabla, podemos calcular fácilmente una *estimación* del valor de la media de estos datos agrupados. Es una estimación porque no utilizamos los 600 datos puntuales de la muestra. De haber usado los datos originales sin agrupar, podríamos haber calculado el valor real de la media, pero sólo después de obtener el promedio de los 600 valores individuales. En aras de la sencillez, debemos sacrificar la precisión.

Estimación de la media

Tabla 3-3
Saldo promedio mensual de 600 cuentas de cheques

Clase (dólares)	Frecuencia
0- 49.99	78
50.00- 99.99	123
100.00-149.99	187
150.00-199.99	82
200.00-249.99	51
250.00-299.99	47
300.00-349.99	13
350.00-399.99	9
400.00-449.99	6
450.00-499.99	4
	600

Cálculo de la media

Para encontrar la media aritmética de datos agrupados, primero calculamos el punto medio de cada clase. Para lograr que los puntos medios queden en cifras cerradas, redondeamos las cantidades. Así, por ejemplo, el punto medio de la primera clase, 24.995, se convierte en 25.00. Después multiplicamos cada punto medio por la frecuencia de las observaciones de dicha clase, sumamos todos los resultados y dividimos esta suma entre el número total de observaciones de la muestra. La fórmula es la siguiente:

Media aritmética de una muestra con datos agrupados

$$\bar{x} = \frac{\Sigma(f \times x)}{n} \qquad [3\text{-}3]$$

donde,

- $\bar{x}$ = media de la muestra
- Σ = símbolo que significa "la suma de"
- f = frecuencia (número de observaciones) de cada clase
- x = punto medio de cada clase en la muestra
- n = número de observaciones en la muestra

En la tabla 3-4 se ilustra cómo calcular la media aritmética de una colección de datos agrupados, utilizando la ecuación 3-3.

Hacemos una suposición

En nuestra muestra de 600 clientes, el saldo mensual promedio de las cuentas de cheques es $142.25. Ésta es la aproximación hecha a partir de la distribución de frecuencias. Observe que, como no conocemos cada uno de los datos puntuales de la muestra, suponemos que todos los valores de una clase son iguales a su punto medio. Nuestros resultados, entonces, son sólo una aproximación del promedio del saldo mensual real.

Codificación

Asignación de códigos o los puntos medios

En aquellas situaciones en que no se tenga disponible una computadora y sea necesario realizar las operaciones aritméticas a mano, podemos simplificar aún más nuestro cálculo de la media de datos agrupados. Mediante una técnica conocida como *codificación*, podemos eliminar el problema de te-

Tabla 3-4

Cálculo de la media aritmética de la muestra con los datos agrupados de la tabla 3-3

Clase (dólares) (1)	Punto medio (x) (2)		Frecuencia (f) (3)		$f \times x$ (3) × (2)
0- 49.99	25.00	×	78	=	1,950
50.00- 99.99	75.00	×	123	=	9,225
100.00-149.99	125.00	×	187	=	23,375
150.00-199.99	175.00	×	82	=	14,350
200.00-249.99	225.00	×	51	=	11,475
250.00-299.99	275.00	×	47	=	12,925
300.00-349.00	325.00	×	13	=	4,225
350.00-399.99	375.00	×	9	=	3,375
400.00-449.99	425.00	×	6	=	2,550
450.00-499.99	475.00	×	4	=	01,900
			$\Sigma f = n =$ **600**		**85,350** ← $\Sigma(f \times x)$

$$\bar{x} = \frac{\Sigma(f \times x)}{n} \qquad [3\text{-}3]$$

$$= \frac{85{,}350}{600}$$

$$= 142.25 \longleftarrow \text{Media de la muestra (dólares)}$$

ner puntos medios muy grandes o inconvenientes. En lugar de utilizar los puntos medios reales en los cálculos, podemos asignar enteros consecutivos de valor pequeño, llamados *códigos*, a cada uno de los puntos medios. El entero cero puede asignarse a cualquier punto medio, pero para que los enteros sean pequeños, asignaremos el cero al punto medio de la *mitad* de la distribución (o el más cercano a la mitad). Entonces podemos asignar enteros negativos a los valores menores que ese punto medio y enteros positivos a los valores más grandes, de la manera siguiente:

Clase	1-5	6-10	11-15	16-20	21-25	26-30	31-35	36-40	41-45
Código (u)	-4	-3	-2	-1	0	1	2	3	4
					↑				
					x_0				

Cálculo de la media de datos agrupados utilizando códigos

Los estadísticos usan x_0 para representar el punto medio al que se asigna el código 0, y u para el punto medio codificado. La siguiente fórmula se utiliza para determinar la media de la muestra mediante códigos:

Media aritmética de la muestra para datos agrupados usando códigos

$$\bar{x} = x_0 + w\,\frac{\Sigma(u \times f)}{n} \qquad [3\text{-}4]$$

donde,

- $\bar{x}$ = media de la muestra
- x_0 = valor del punto medio al que se asignó el código 0
- w = ancho numérico del intervalo de clase
- u = código asignado a cada punto medio de clase
- f = frecuencia o número de observaciones de cada clase
- n = número total de observaciones de la muestra

Tenga en mente que $\Sigma(u \times f)$ simplemente significa que 1) multiplicamos u por f para cada clase en la distribución de frecuencias, y 2) sumamos todos estos productos. La tabla 3-5 ilustra cómo codi-

Tabla 3-5

Caída anual de nieve en Harlan, Kentucky

Clase (1)	Punto medio (x) (2)	Código (u) (3)		Frecuencia (f) (4)		$u \times f$ (3) × (4)
0- 7	3.5	-2	×	2	=	-4
8-15	11.5	-1	×	6	=	-6
16-23	19.5←x_0	0	×	3	=	0
24-31	27.5	1	×	5	=	5
32-39	35.5	2	×	2	=	4
40-47	43.5	3	×	2	=	6
				$\Sigma f = n =$ **20**		5←$\Sigma(u \times f)$

$$\bar{x} = x_0 + w\,\frac{\Sigma(u \times f)}{n} \qquad [3\text{-}4]$$

$$= 19.5 + 8\left(\frac{5}{20}\right)$$

$$= 19.5 + 2$$

$$= 21.5 \longleftarrow \text{Caída de nieve anual promedio}$$

ficar los puntos medios y encontrar la media de la muestra de la caída anual de nieve (en pulgadas) durante 20 años en Harlan, Kentucky.

Ventajas y desventajas de la media aritmética

Ventajas de la media

La media aritmética, como un solo número que representa a un conjunto de datos completo, tiene importantes ventajas. Primero, se trata de un concepto familiar para la mayoría de las personas y es intuitivamente claro. Segundo, cada conjunto de datos tiene una media; es una medida que puede calcularse y es única debido a que cada conjunto de datos posee una y sólo una media. Por último, la media es útil para llevar a cabo procedimientos estadísticos como la comparación de medias de varios conjuntos de datos (procedimiento que se estudiará en el capítulo 9).

Tres desventajas de la media

Sin embargo, como cualquier medida estadística, la media aritmética tiene desventajas que debemos conocer. **Primero**, aunque la media es confiable en cuanto a que toma en cuenta todos los valores del conjunto de datos, puede verse afectada por valores extremos que no son representativos del resto de los datos. Observe que si los siete miembros de un equipo de atletismo tienen las marcas de tiempo que se muestran en la tabla 3-6 para cierta carrera, el tiempo medio es:

$$\mu = \frac{\Sigma x}{N} \qquad [3\text{-}1]$$

$$= \frac{4.2 + 4.3 + 4.7 + 4.8 + 5.0 + 5.1 + 9.0}{7}$$

$$= \frac{37.1}{7}$$

$$= 5.3 \text{ minutos} \longleftarrow \text{Media de la población}$$

Sin embargo, si calculamos el tiempo medio para los primeros seis corredores y excluimos el valor de 9.0 minutos, la respuesta aproximada es 4.7 minutos. El *valor extremo* 9.0 distorsiona el valor que obtenemos para la media. Sería más representativo calcular la media *sin* incluir el valor extremo.

Un **segundo** problema con la media es el mismo que encontramos con los 600 saldos de cuentas de cheques. Resulta tedioso calcular la media debido a que *utilizamos* cada uno de los datos en nuestro cálculo (a menos, desde luego, que usemos el método corto que consiste en utilizar datos agrupados para determinar la media aproximada).

La **tercera** desventaja es que somos incapaces de calcular la media para un conjunto de datos que tiene clases de extremo abierto en la parte inferior o superior de la escala. Suponga que los datos de la tabla 3-6 se clasifican en la distribución de frecuencias de la tabla 3-7. No podemos calcular un valor para la media de estos datos debido a la clase de extremo abierto "5.4 o más". No tenemos forma de saber si el valor de la observación de esta clase es 5.4, cercano a 5.4 o mucho mayor que 5.4.

Tabla 3-6
Tiempos de los integrantes de un equipo de atletismo en una carrera de 1 milla

Integrante	1	2	3	4	5	6	7
Tiempo en minutos	4.2	4.3	4.7	4.8	5.0	5.1	9.0

Tabla 3-7
Tiempos de los integrantes de un equipo de atletismo en una carrera de 1 milla

Clase en minutos	4.2-4.5	4.6-4.9	5.0-5.3	5.4 o más
Frecuencia	2	2	2	1

SUGERENCIAS Y SUPOSICIONES La media (o promedio) *puede* ser una excelente medida de tendencia central (la manera en que se agrupan los datos alrededor del punto medio de una distribución); pero a menos que la media sea en verdad representativa de los datos con los que se calculó, estaríamos violando una suposición importante. Advertencia: si existen valores muy altos o muy bajos notoriamente distintos a la mayoría de los datos, la media *no* es representativa. Por fortuna, existen medidas que se pueden calcular que no tienen este defecto. Una sugerencia útil al elegir qué medidas calcular es observar los datos.

Ejercicios 3-2

Ejercicios de autoevaluación

EA 3-1 La siguiente distribución de frecuencias representa los pesos en libras de una muestra de paquetes transportados el mes pasado por una pequeña compañía de carga aérea.

Clase	Frecuencia	Clase	Frecuencia
10.0-10.9	1	15.0-15.9	11
11.0-11.9	4	16.0-16.9	8
12.0-12.9	6	17.0-17.9	7
13.0-13.9	8	18.0-18.9	6
14.0-14.9	12	19.0-19.9	2

a) Calcule la media de la muestra con la ecuación 3-3.
b) Calcule la media de la muestra usando el método de códigos (ecuación 3-4) con 0 asignado a la cuarta clase.
c) Repita el inciso b) con 0 asignado a la sexta clase.
d) Explique por qué sus repuestas a los incisos b) y c) son iguales.

EA 3-2 La Davis Furniture Company tiene un acuerdo de crédito revolvente con el First National Bank. El préstamo mostró los siguiente saldos de fin de mes durante el año pasado

Ene.	$121,300	Abr.	$72,800	Jul.	$58,700	Oct.	$52,800
Feb.	$112,300	May.	$72,800	Ago.	$61,100	Nov.	$49,200
Mar.	$ 72,800	Jun.	$57,300	Sep.	$50,400	Dic.	$46,100

La compañía puede obtener una tasa de interés menor si su saldo mensual promedio es mayor que $65,000. ¿Califica para esa tasa de interés menor?

Aplicaciones

■ **3-6** El Child-Care Community Nursery es elegible para recibir recursos de un fondo especial de servicios sociales del estado, siempre y cuando la edad promedio de sus niños esté por debajo de los nueve años. Si los datos que se presentan a continuación representan la edad de los niños que acuden normalmente al centro, ¿calificará éste para el apoyo del fondo?

8 5 9 10 9 12 7 12 13 7 8

■ **3-7** El Child-Care Community Nursery puede continuar recibiendo el apoyo económico de servicios sociales del estado siempre y cuando el promedio del ingreso anual de las familias cuyos niños asisten al centro sea menor que $12,500. Los ingresos familiares de los niños del centro son:

$14,500	$15,600	$12,500	$8,600	$ 7,800	
$ 6,500	$ 5,900	$10,200	$8,800	$14,300	$13,900

a) ¿El centro en cuestión sigue calificando para recibir apoyo?

b) Si la respuesta del inciso a) es no, ¿cuánto debe disminuir el ingreso familiar promedio para que el centro califique?

c) Si la respuesta del inciso a) es sí, ¿cuánto puede aumentar el ingreso familiar promedio y todavía seguir calificando?

■ **3-8** Los siguientes datos representan las edades de los pacientes admitidos en un pequeño hospital el día 28 de febrero de 1996:

85	75	66	43	40
88	80	56	56	67
89	83	65	53	75
87	83	52	44	48

a) Construya una distribución de frecuencias con clases 40-49, 50-59, etcétera.

b) Calcule la media de la muestra a partir de la distribución de frecuencias.

c) Calcule la media de la muestra a partir de los datos sin procesar.

d) Compare los incisos b) y c) y comente su respuesta.

■ **3-9** La siguiente distribución de frecuencias representa el tiempo en segundos que los cajeros de BullsEye Discount Store necesitaron para servir a una muestra de clientes en diciembre de 1996.

Tiempo (en segundos)	Frecuencia
20- 29	6
30- 39	16
40- 49	21
50- 59	29
60- 69	25
70- 79	22
80- 89	11
90- 99	7
100-109	4
110-119	0
120-129	2

a) Calcule la media de la muestra con la ecuación 3-3.

b) Calcule la media de la muestra usando el método de códigos (ecuación 3-4) con 0 asignado a la clase 70-79.

■ **3-10** El dueño de Pets'R Us está interesado en construir una nueva tienda. La construirá si el número promedio de animales vendidos durante los primeros 6 meses de 1995 es al menos 300 y si el promedio mensual global del año es al menos 285. Los datos para 1995 son los siguientes:

Ene.	Feb.	Mar.	Abr.	May.	Jun.	Jul.	Ago.	Sep.	Oct.	Nov.	Dic.
234	216	195	400	315	274	302	291	275	300	375	450

¿Qué decisión toma el dueño y por qué?

■ **3-11** Un fabricante de cosméticos adquirió una máquina para llenar botellas de perfume de 3 onzas. Para probar la precisión del volumen depositado en cada botella, hizo una corrida de prueba con 18 recipientes. Los volúmenes resultantes (en onzas) de la prueba fueron los siguientes:

3.02	2.89	2.92	2.84	2.90	2.97	2.95	2.94	2.93
3.01	2.97	2.95	2.90	2.94	2.96	2.99	2.99	2.97

La compañía no suele recalibrar la máquina para este perfume si el volumen de llenado de las 3 onzas difiere en 0.04 onzas o menos. ¿Deberá recalibrarla?

■ **3-12** El gerente de producción de la imprenta Hinton desea determinar el tiempo promedio necesario para fotografiar una placa de impresión. Utilizando un cronómetro y observando a los operadores, registra los tiempos siguientes (en segundos):

20.4	20.0	22.2	23.8	21.3	25.1	21.2	22.9	28.2	24.3
22.0	24.7	25.7	24.9	22.7	24.4	24.3	23.6	23.2	21.0

Un tiempo promedio por placa menor a los 23.0 segundos indica una productividad satisfactoria. ¿Debe estar preocupado el gerente de producción?

■ **3-13** La National Tire Company tiene sus fondos de reserva en una inversión a corto plazo. El saldo diario (en millones de dólares) de la cuenta de inversión durante 2 semanas es el siguiente:

Semana 1	$1.973	$1.970	$1.972	$1.975	$1.976
Semana 2	1.969	1.892	1.893	1.887	1.895

¿Cuál es la cantidad promedio (media) invertida durante

a) la primera semana?
b) la segunda semana?
c) el periodo de 2 semanas?
d) Un saldo promedio durante las 2 semanas mayor que $1.970 millones calificaría a National para obtener tasas de interés más altas. ¿Califica?
e) Si la respuesta del inciso c) es menor que $1.970 millones, ¿cuánto tendría que aumentar la cantidad invertida el último día para que la compañía obtuviera las tasas de interés más altas?
f) Si la repuesta del inciso c) es mayor que $1.970 millones, ¿cuánto podría el tesorero de la compañía retirar el último día de los fondos de reserva, de manera que todavía calificara para las tasas de interés altas?

■ **3-14** M.T. Smith recorre el este de Estados Unidos como representante de ventas del editor de un libro de texto. Recibe una comisión proporcional al volumen de las ventas que haga. Sus ganancias trimestrales en dólares durante los últimos tres años son las siguientes:

	1er. trimestre	2do. trimestre	3er. trimestre	4to. trimestre
Año 1	$10,000	$ 5,000	$25,000	$15,000
Año 2	20,000	10,000	20,000	10,000
Año 3	30,000	15,000	45,000	50,000

a) Calcule por separado las ganancias promedio de la representante en cada uno de los cuatro trimestres.
b) Calcule por separado las ganancias trimestrales promedio en cada uno de los tres años.
c) Muestre que la media de las cuatro cantidades obtenida en el inciso a) es igual a la media de las tres cantidades que obtuvo en el inciso b). Además, muestre que estas dos cantidades son iguales a la media de los 12 números que se presentan en la tabla. (Ésta es la ganancia promedio trimestral que obtuvo la señorita Smith durante un periodo de tres años.)

■ **3-15** Lillian Tyson ha sido, durante diez años, la presidenta del comité organizador de la biblioteca municipal. Afirma que durante su cargo ha administrado el presupuesto para el mantenimiento de la biblioteca ambulante del municipio mejor que su antecesor. A continuación presentamos los datos relativos al mantenimiento de la biblioteca ambulante durante quince años, en dólares:

Año	Presupuesto	Año	Presupuesto	Año	Presupuesto
1992	$30,000	1987	$24,000	1982	$30,000
1991	28,000	1986	19,000	1981	20,000
1990	25,000	1985	21,000	1980	15,000
1989	27,000	1984	22,000	1979	10,000
1988	26,000	1983	24,000	1978	9,000

a) Calcule el presupuesto promedio anual para los últimos 5 años (1988-1992).
b) Calcule el presupuesto promedio anual para los primeros 5 años de gestión (1983-1987).
c) Calcule el presupuesto promedio anual para los 5 años anteriores a su elección (1978-1982).
d) Basándose en los resultados de los incisos a), b) y c), ¿podría concluir que ha habido una tendencia a aumentar o a disminuir en el presupuesto anual? ¿La presidenta actual ha ahorrado dinero al municipio?

Soluciones a los ejercicios de autoevaluación

EA 3-1

		(a)		(b)		(c)	
Clase	Frecuencia (f)	Punto medio (x)	$f \times x$	Código u	$u \times f$	Código u	$u \times f$
10.0-10.9	1	10.5	10.5	−3	−3	−5	−5
11.0-11.9	4	11.5	46.0	−2	−8	−4	−16
12.0-12.9	6	12.5	75.0	−1	−6	−3	−18
13.0-13.9	8	13.5	108.0	0	0	−2	−16
14.0-14.9	12	14.5	174.0	1	12	−1	−12
15.0-15.9	11	15.5	170.5	2	22	0	0
16.0-16.9	8	16.5	132.0	3	24	1	8
17.0-17.9	7	17.5	122.5	4	28	2	14
18.0-18.9	6	18.5	111.0	5	30	3	18
19.0-19.9	2	19.5	39.0	6	12	4	8
	65		**988.5**		**111**		**−19**

a) $\bar{x} = \dfrac{\Sigma(f \times x)}{n} = \dfrac{988.5}{65} = 15.2077$ libras

b) $\bar{x} = x_0 + w\,\dfrac{\Sigma(u \times f)}{n} = 13.5 + \dfrac{1.0(111)}{65} = 15.2077$ libras

c) $\bar{x} = x_0 + w\,\dfrac{\Sigma(u \times f)}{n} = 15.5 + \dfrac{1.0(-19)}{65} = 15.2077$ libras

d) Al mover la clase con el código 0 asignado k clases hacia arriba, se sustituye x_0 por $x_0 + kw$ y se cambia cada código de u a $u - k$. Pero como

$$\bar{x}_b = x_0 + w\,\frac{\Sigma(u \times f)}{n} = (x_0 + kw) - kw + w\,\frac{\Sigma(u \times f)}{n}$$

$$= (x_0 + kw) + w\,\frac{\Sigma(u - k)f}{n} = \bar{x}_c$$

se ve que no importa a qué clase se asigne el código 0.

EA 3-2 $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{827{,}600}{12} = \$68{,}967$

Dado que esto excede \$65,000, la compañía califica para las tasas de interés reducidas.

3.3 Una segunda medida de tendencia central: la media ponderada

Una media ponderada

La media ponderada nos permite calcular un promedio que toma en cuenta la *importancia* de cada valor con respecto al total. Considere, por ejemplo, la compañía cuyos datos presentamos en la tabla 3-8; ésta utiliza tres niveles de trabajo —no calificado, semicalificado y calificado— para la producción de dos de sus productos finales. La compañía desea saber el promedio del costo de trabajo por hora para cada uno de los productos.

Tabla 3-8

Mano de obra por proceso de manufactura

Nivel de mano de obra	Salario por hora en dólares (x)	Horas de mano de obra por unidad producida	
		Producto 1	Producto 2
No calificado	\$5.00	1	4
Semicalificado	7.00	2	3
Calificado	9.00	5	3

Un simple promedio aritmético de los salarios pagados sería:

$$\bar{x} = \frac{\Sigma x}{n} \qquad [3\text{-}2]$$

$$= \frac{\$5 + \$7 + \$9}{3}$$

$$= \frac{\$21}{3}$$

$$= \$7.00/\text{hora}$$

En este caso la media aritmética es incorrecta

Usando esta tasa promedio podríamos calcular el costo del trabajo invertido en una unidad del producto 1 como $\$7(1 + 2 + 5) = \$56$, y el de una unidad del producto 2 como $\$7(4 + 3 + 3) = \$70$. Pero estos promedios son incorrectos.

Para que nuestros cálculos sean correctos, las respuestas deben tomar en cuenta que se utilizan diferentes niveles de mano de obra. Podemos determinar los promedios correctos de la siguiente manera. Para el producto 1, el costo total del trabajo por unidad es $(\$5 \times 1) + (\$7 \times 2) + (\$9 \times 5) = \$64$, y como se invierten ocho horas de trabajo, el costo promedio de mano de obra por hora es $\$64/8 = \$8.00$. Para el producto 2, el costo total del trabajo por unidad es $(\$5 \times 4) + (\$7 \times 3) + (\$9 \times 3) = \$68$, para un costo promedio de mano de obra por hora de $\$68/10 = \$6.80$.

La respuesta correcta es la media ponderada

Otra forma de calcular el costo promedio por hora para los dos productos consiste en tomar un *promedio ponderado* del costo de los tres niveles de mano de obra. Para hacerlo, *ponderamos* el salario por hora de cada nivel mediante la proporción de la mano de obra total requerida para fabricar el producto. Una unidad del producto 1, por ejemplo, requiere 8 horas de trabajo. De este tiempo, $^1/_8$ es de mano obra no calificada, $^2/_8$ de mano de obra semicalificada y $^5/_8$ de trabajo calificado. Si utilizamos estas fracciones como las ponderaciones (o los pesos), entonces una hora de trabajo en el producto 1 cuesta en promedio:

$$\left(\frac{1}{8} \times \$5\right) + \left(\frac{2}{8} \times \$7\right) + \left(\frac{5}{8} \times \$9\right) = \$8.00/\text{hora}$$

De manera análoga, una unidad del producto 2 requiere 10 horas de mano de obra; de las cuales $^4/_{10}$ son de trabajo no calificado, $^3/_{10}$ de trabajo semicalificado y $^3/_{10}$ de trabajo calificado. Utilizando estas fracciones como ponderaciones o pesos, una hora de mano de obra en el producto 2 cuesta:

Cálculo de la media ponderada

$$\left(\frac{4}{10} \times \$5\right) + \left(\frac{3}{10} \times \$7\right) + \left(\frac{3}{10} \times \$9\right) = \$6.80/\text{hora}$$

Así, vemos que los promedios ponderados dan el valor correcto para los costos promedio por hora de mano de obra de los dos productos, ya que **consideran las diferentes cantidades de cada nivel de mano de obra que requieren los productos**.

Con símbolos, la fórmula para calcular el promedio ponderado es:

Media ponderada

$$\bar{x}_w = \frac{\Sigma(w \times x)}{\Sigma w} \qquad [3\text{-}5]$$

donde,

- $\bar{x}_w$ = símbolo para la media ponderada*
- w = peso asignado a cada observación ($^1/_8$, $^2/_8$ y $^5/_8$ para el producto 1, y $^4/_{10}$, $^3/_{10}$ y $^3/_{10}$ para el producto 2 del ejemplo)
- $\Sigma(w \times x)$ = la suma de los productos de la ponderación de cada elemento por el elemento correspondiente
- Σw = suma de todas las ponderaciones

Si aplicamos la ecuación 3-5 al producto 1 de nuestro ejemplo de costo de mano de obra, encontramos que

$$\bar{x}_w = \frac{\Sigma(w \times x)}{\Sigma w} \qquad [3\text{-}5]$$

$$= \frac{\left(\frac{1}{8} \times \$5\right) + \left(\frac{2}{8} \times \$7\right) + \left(\frac{5}{8} \times \$9\right)}{\frac{1}{8} + \frac{2}{8} + \frac{5}{8}}$$

$$= \frac{\$8}{1}$$

$$= \$8.00/\text{hora}$$

Media aritmética de datos agrupados: la media ponderada

Observe que la ecuación 3-5 establece de una manera más formal algo que ya habíamos hecho. Cuando calculamos la media aritmética de datos agrupados, en realidad encontramos una media aritmética ponderada, utilizando los puntos medios como valores de x y las frecuencias de cada clase como pesos (o ponderaciones). Dividimos este producto entre la suma de todas las frecuencias, que es igual a la división entre la suma de todos los pesos.

De manera análoga, *cualquier* media calculada a partir de todos los valores de un conjunto de datos, de acuerdo con la ecuación 3-1 o 3-2 es, en realidad, el promedio ponderado de los componentes del conjunto de datos. Desde luego, la naturaleza de tales componentes determina qué es lo que la media está midiendo. En una fábrica, por ejemplo, podemos determinar la media ponderada de todos los tipos de salarios (no calificado, semicalificado y calificado) o salarios de trabajadores hombres y mujeres o de trabajadores sindicalizados y no sindicalizados.

SUGERENCIAS Y SUPOSICIONES

Debe hacerse la distinción entre *valores diferentes* y *observaciones individuales* en un conjunto de datos, ya que varias observaciones pueden tener el mismo valor. Si los valores ocurren con frecuencias diferentes, la media aritmética de los *valores* (comparada con la media aritmética de las *observaciones*) tal vez no sea una medida de tendencia central exacta. En esos casos, es necesario usar la media ponderada de los valores. Si se utiliza un valor promedio para tomar una decisión, pregunte cómo se calculó. Si los valores de la muestra no aparecen con la misma frecuencia, insista en que la base correcta para la toma de decisiones es la media ponderada.

*El símbolo $\bar{x}_w$ se lee *x barra sub w*. La letra w se conoce como subíndice y sirve para recordar que no se trata de una media ordinaria, sino de una media ponderada, de acuerdo con la importancia relativa de los valores de x.

Ejercicios 3.3

Ejercicios de autoevaluación

EA 3-3 La tienda Dave's Giveaway tiene un aviso: "Si nuestros precios promedio no son iguales o menores que los de otros, usted se lo lleva gratis." Uno de los clientes de Dave's fue a la tienda un día y puso sobre el mostrador las notas de venta de seis artículos que compró a un competidor por un precio promedio menor que el de Dave's. Los artículos costaron (en dólares)

$1.29 $2.97 $3.49 $5.00 $7.50 $10.95

Los precios de Dave's de los mismos seis artículos son $2.35, $2.89, $3.19, $4.98, $7.59 y $11.50. Dave's le explicó al cliente: "Mi aviso se refiere a un promedio ponderado de estos artículos, nuestro promedio es menor porque nuestras ventas de estos artículos han sido:

7 9 12 8 6 3

¿Está Dave's buscando un problema o resolviéndolo al hablar de promedios ponderados?

EA 3-4 La Bennett Distribution Company, una subsidiaria de un importante fabricante de electrodomésticos, desea pronosticar las ventas regionales para el año próximo. Se espera que la sucursal de la costa del Atlántico, con ventas actuales de $193.8 millones, logre un crecimiento en las ventas del 7.25%; se espera que la sucursal del Medio Oeste, con ventas actuales de $79.3 millones, tenga un incremento del 8.20%, y que la sucursal de la costa del Pacífico, con ventas actuales de $57.5 millones, aumente sus ventas 7.15%. ¿Cuál es la tasa promedio de crecimiento pronosticado en las ventas para el próximo año?

Aplicaciones

■ **3-16** Un profesor decide utilizar un promedio ponderado para obtener las calificaciones finales de los estudiantes que acuden a su seminario. El promedio de tareas tendrá un valor del 20% de la calificación del estudiante; el examen semestral, 25%; el examen final, 35%; el artículo de fin de semestre, 10%, y los exámenes parciales, 10%. A partir de los datos siguientes, calcule el promedio final para los cinco estudiantes del seminario.

Estudiante	Tareas	Parciales	Artículo	Ex. semestral	Ex. final
1	85	89	94	87	90
2	78	84	88	91	92
3	94	88	93	86	89
4	82	79	88	84	93
5	95	90	92	82	88

■ **3-17** Jim's Videotaping Service hizo un pedido de cintas VHS. Jim ordenó 6 cajas de High-Grade, 4 cajas de Performance High-Grade, 8 cajas de Standard, 3 cajas de High Standard y 1 caja de Low Grade. Cada caja contiene 24 cintas. Suponga que los costos por caja son: High-Grade, $28; Performance High-Grade, $36; Standard, $16; High Standard, $18, y Low, $6.

a) ¿Cuál es el costo promedio por caja?
b) ¿Cuál es el costo promedio por cinta?
c) Suponga que Jim's piensa vender cualquier cinta por $1.25, ¿sería esto un buen negocio para Jim's?
d) ¿Cómo cambiaría su respuesta a los incisos a) a c) si hubiera 48 cintas por caja?

■ **3-18** La mueblería Keyes publicó seis anuncios en los periódicos locales durante el mes de diciembre. Como resultado se obtuvo la siguiente distribución de frecuencias:

NÚMERO DE VECES QUE UN LECTOR VIO EL ANUNCIO DURANTE DICIEMBRE	0	1	2	3	4	5	6
FRECUENCIA	897	1,082	1,325	814	307	253	198

¿Cuál es el número promedio de veces que un lector vio un anuncio de la mueblería Keyes durante diciembre?

■ **3-19** La Nelson Window Company tiene plantas de manufactura en cinco ciudades de Estados Unidos: Orlando, Minneapolis, Dallas, Pittsburgh y Seattle. Se elaboró el pronóstico de producción para el próximo año. La división de Orlando, con una producción anual de 72 millones de ventanas, pronostica un incremento del 11.5%. La división de Pittsbrugh, con producción anual de 62 millones, debe crecer 6.4%. La división de Seattle, cuya producción anual es 48 millones, también debe crecer 6.4%. Se espera que las divisiones de Minneapolis y Dallas, con producciones respectivas de 89 y 94 millones cada año, tengan disminuciones del 9.7 y 18.2%, respectivamente. ¿Cuál es la tasa promedio de cambio en producción para la Nelson Window Company durante el año próximo?

■ **3-20** El Servicio Postal de Estados Unidos maneja siete tipos básicos de cartas y tarjetas postales: tercera clase, segunda clase, primera clase, correo aéreo, entrega especial, correo registrado y correo certificado. El volumen de envíos durante 1977 se da en la siguiente tabla:

Tipo de correo	Onzas enviadas (en millones)	Precio por cada onza
Tercera clase	16,400	$0.05
Segunda clase	24,100	0.08
Primera clase	77,600	0.13
Aéreo	1,900	0.17
Entrega especial	1,300	0.35
Registrado	750	0.40
Certificado	800	0.45

¿Cuál es el ingreso promedio anual por cada onza de la prestación del servicio?

■ **3-21** Matthews, Young y Asociados, un despacho de asesoría financiera y administrativa, tiene cuatro tipos de profesionales entre su personal: asesores financieros, asociados principales, personal de campo y personal de oficina. Las tasas promedio que se cobran a los clientes por el desempeño de cada una de estas categorías profesionales son 75 dólares/hora, 40 dólares/hora, 30 dólares/hora y 15 dólares/hora, respectivamente. Los registros de la firma indican el siguiente número de horas cobradas el año anterior en cada categoría: 8,000, 14,000, 24,000 y 35,000, respectivamente. Si Mathews, Young y Asociados intenta formular una tasa de cobro promedio para estimar cuánto debe cobrar a los clientes en el año siguiente, ¿qué sugeriría que hiciera y cuál cree que sería una tasa apropiada?

Soluciones a los ejercicios de autoevaluación

EA **3-3** Con los promedios no ponderados, se obtiene

$$\bar{x}_c = \frac{\Sigma x}{n} = \frac{31.20}{6} = \$5.20 \text{ en la competencia}$$

$$\bar{x}_D = \frac{31.50}{6} = 5.25 \text{ en la tienda Dave}$$

Con los promedios ponderados, se obtiene

$$\bar{x}_c = \frac{\Sigma(w \times x)}{\Sigma w}$$

$$= \frac{7(1.29) + 9(2.97) + 12(3.49) + 8(5.00) + 6(7.50) + 3(10.95)}{7 + 9 + 12 + 8 + 6 + 3}$$

$$= \frac{195.49}{45} = \$4.344 \text{ en la competencia}$$

$$\bar{x}_D = \frac{7(1.35) + 9(2.89) + 12(3.19) + 8(4.98) + 6(7.59) + 3(11.50)}{7 + 9 + 12 + 8 + 6 + 3}$$

$$= \frac{193.62}{45} = \$4.303 \text{ en la tienda Dave}$$

Aunque en términos técnicos Dave está en lo correcto, la palabra *promedio* en el uso popular es equivalente al *promedio no ponderado* del uso técnico y es seguro que el cliente típico se molestará con la afirmación de Dave (entienda o no el matiz técnico).

EA 3-4

$$\bar{x}_w = \frac{\Sigma(w \times x)}{\Sigma w} = \frac{193.8(7.25) + 79.3(8.20) + 57.5(7.15)}{193.8 + 79.3 + 57.5}$$

$$= \frac{2466.435}{330.6} = 7.46\%$$

3.4 Una tercera medida de tendencia central: la media geométrica

Búsqueda de la tasa de crecimiento: la media geométrica

Algunas veces, cuando trabajamos con cantidades que cambian en cierto periodo, necesitamos conocer una tasa promedio de cambio, como la tasa de crecimiento promedio en un periodo de varios años. En tales casos, la media aritmética simple resulta inapropiada, pues proporciona resultados equivocados. Lo que debemos encontrar es la *media geométrica*, llamada simplemente la M.G.

Considere, por ejemplo, el crecimiento de una cuenta de ahorros. Suponga que inicialmente depositamos $100 y dejamos que acumule intereses a diferentes tasas durante cinco años. El crecimiento se resume en la tabla 3-9.

La entrada con el encabezado "factor de crecimiento" es igual a:

$$1 + \frac{\text{tasa de interés}}{100}$$

El factor de crecimiento es la cantidad por la que multiplicamos los ahorros al inicio del año para obtener el saldo al final del mismo. El factor de crecimiento considerado como la media aritmética simple sería $(1.07 + 1.08 + 1.10 + 1.12 + 1.18)/5 = 1.11$, que corresponde a una tasa de interés promedio del 11% anual. Sin embargo, si el banco diera intereses a una tasa constante del 11% anual, un depósito de $100 crecería en cinco años a:

En este caso, la tasa de crecimiento tomada como la media aritmética es incorrecta

$$\$100 \times 1.11 \times 1.11 \times 1.11 \times 1.11 \times 1.11 = \$168.51$$

En la tabla 3-9 se muestra que la cifra real es sólo $168.00. Así, el factor de crecimiento promedio correcto debe ser ligeramente menor a 1.11.

Cálculo de la media geométrica

Para encontrar el factor de crecimiento promedio correcto podemos multiplicar los factores de crecimiento de los cinco años y luego obtener la raíz quinta del producto (número que al multiplicarse cuatro veces por sí mismo da como resultado el producto inicial). El resultado es el *factor de crecimiento como media geométrica*, que es el promedio adecuado que debemos utilizar. La fórmula para encontrar la media geométrica de una serie de números es:

Media geométrica

Número de valores x

$$\text{M.G.} = \sqrt[n]{\text{producto de todos los valores } x} \qquad [3\text{-}6]$$

Si aplicamos esta ecuación a nuestro problema de la cuenta de ahorros, podemos determinar que 1.1093 es el factor de crecimiento promedio correcto.

$$\text{M.G.} = \sqrt[n]{\text{Producto de todos los valores } x} \qquad [3\text{-}6]$$

$$= \sqrt[5]{1.07 \times 1.08 \times 1.10 \times 1.12 \times 1.18}$$

Tabla 3-9

Crecimiento de un depósito de $100 en una cuenta de ahorros

Año	Tasa de interés (porcentaje)	Factor de crecimiento	Ahorros al final de año (dólares)
1	7%	1.07	$107.00
2	8	1.08	115.56
3	10	1.10	127.12
4	12	1.12	142.37
5	18	1.18	168.00

$$= \sqrt[5]{1.679965}$$

$$= 1.1093 \longleftarrow \text{Factor de crecimiento promedio (media geométrica de los 5 factores de crecimiento)}$$

Advertencia: utilice la media apropiada

Observe que la tasa de interés promedio correcta del 10.93% anual obtenida con la media geométrica está muy cerca de la tasa promedio incorrecta del 11% anual obtenida con la media aritmética. Esto se debe a que las tasas de interés son relativamente pequeñas. Sin embargo, tenga cuidado de no verse tentado a utilizar la media aritmética en lugar de la geométrica, que es más complicada. El siguiente ejemplo nos muestra por qué.

En las economías con un alto índice de inflación, los bancos deben pagar altas tasas de interés para atraer a los ahorradores. Suponga que en un periodo de cinco años en un régimen económico con un muy alto índice de inflación, los bancos pagan tasas de interés anual de 100, 200, 250, 300 y 400%, que corresponde a un factor de crecimiento de 2, 3, 3.5, 4 y 5. (Calculamos estos factores de crecimiento del mismo modo que en la tabla 3-9.)

En cinco años, un depósito inicial de $100 crecerá a $100 × 2 × 3 × 3.5 × 4 × 5 = $42,000. El factor de crecimiento como media aritmética es de (2 + 3 + 3.5 + 4 + 5)/5 = 3.5. Esto corresponde a una tasa de interés promedio anual del 250%. Sin embargo, si el banco en realidad pagara intereses a una tasa constante de 250 anual, entonces $100 crecerían a $52,521.88 en cinco años:

$$\$100 \times 3.5 \times 3.5 \times 3.5 \times 3.5 \times 3.5 = \$52{,}521.88$$

Este resultado excede al resultado real de $42,000 en más de $10,500, un error considerable.

Utilicemos la fórmula para obtener la media geométrica de una serie de números para determinar el factor de crecimiento correcto:

$$\text{M.G.} = \sqrt[n]{\text{producto de todos los valores } x} \qquad [3\text{-}6]$$

$$= \sqrt[5]{2 \times 3 \times 3.5 \times 4 \times 5}$$

$$= \sqrt[5]{420}$$

$$= 3.347 \longleftarrow \text{Factor de crecimiento promedio}$$

Este factor de crecimiento corresponde a una tasa de interés promedio del 235% anual. En este caso, el uso de la media apropiada *conduce* a una diferencia significativa.

SUGERENCIAS Y SUPOSICIONES

Se usa la media geométrica para mostrar los efectos multiplicativos en el tiempo de los cálculos del interés compuesto y la inflación. En ciertas situaciones, las respuestas obtenidas con la media aritmética no difieren mucho de las correspondientes a la media geométrica, pero incluso diferencias pequeñas pueden generar malas decisiones. Una buena sugerencia de trabajo es usar la media geométrica siempre que se desee calcular el cambio porcentual promedio en el tiempo para algunas variables. Cuando vea el valor del incremento promedio en la inflación, por ejemplo, pregunte si se trata de la media geométrica y tenga cuidado si no lo es, pues se está manejando un valor incorrecto.

Ejercicios 3.4

Ejercicios de autoevaluación

EA 3-5 El crecimiento en el gasto por deudores morosos de Johnston Office Supply Company durante los últimos años es el siguiente. Calcule el incremento promedio porcentual del gasto por deudores morosos durante ese periodo. Si esta tasa continúa, estime el incremento porcentual para 1977 respecto a 1995.

1989	1990	1991	1992	1993	1994	1995
0.11	0.09	0.075	0.08	0.095	0.108	0.120

EA 3-6 Las tiendas Realistic Stereo etiquetan su mercancía 35% arriba del costo de su última adición al inventario. Hasta hace 4 meses, la grabadora Dynamic 400-S VHS costaba $300. Durante los últimos 4 meses Realistic recibió 4 embarques mensuales de esta grabadora con los siguientes costos unitarios: $275, $250, $240 y $225. ¿A qué tasa promedio mensual ha disminuido el precio de venta de Realistic en estos 4 meses?

Aplicaciones

■ **3-22** Hayes Textiles ha mostrado los siguientes aumentos porcentuales en su valor neto durante los últimos 5 años:

1992	1993	1994	1995	1996
5%	10.5%	9.0%	6.0%	7.5%

¿Cuál es el aumento porcentual promedio del valor neto en el periodo de 5 años?

■ **3-23** MacroSwift, el gigante de software en Estados Unidos, ha publicado un incremento en su valor neto durante 7 de los últimos 9 años. Calcule el cambio porcentual promedio en el valor neto durante este periodo. Suponga condiciones similares en los 3 años siguientes y estime el cambio porcentual para 1998 respecto a 1996.

1988	1989	1990	1991	1992	1993	1994	1995	1996
0.11	0.09	0.07	0.08	−0.04	0.14	0.11	−0.03	0.06

■ **3-24** La compañía Birch, fabricante de tableros de circuitos eléctricos, ha producido el siguiente número de unidades en los últimos cinco años:

1992	1993	1994	1995	1996
12,500	13,250	14,310	15,741	17,630

Calcule el aumento porcentual promedio de unidades producidas en este periodo, y utilice el resultado para estimar la producción en 1999.

■ **3-25** Bob Headen desea calcular el factor de crecimiento promedio de su tienda de aparatos de sonido en los últimos 6 años; utilizando una media geométrica, llega a un resultado de 1.24. Los factores de crecimiento individuales de los últimos 5 años fueron 1.19, 1.35, 1.23, 1.19 y 1.30, pero Bob perdió los registros del sexto año después de haber calculado la media. ¿Cuál era ese factor de crecimiento?

■ **3-26** En un periodo de 3 semanas, el dueño de una tienda adquirió $120 de cubierta de acrílico para forrar sus nuevos mostradores; hizo la adquisición en tres compras de $40 cada una. La primera compra fue a $1.00 el pie cuadrado; la segunda, a $1.10 y la tercera, a $1.15. ¿Cuál fue la tasa de crecimiento promedio semanal en el precio por pie cuadrado que pagó por la cubierta?

■ **3-27** Lisa's Quick Stop atrae a sus clientes con la venta de leche a un precio 2% menor que la tienda de abarrotes más grande del pueblo. Los siguientes son los precios de un galón de leche durante un periodo de 2 meses. ¿Cuál es el cambio porcentual promedio del precio en la tienda de Lisa?

Semana 1	Semana 2	Semana 3	Semana 4	Semana 5	Semana 6	Semana 7	Semana 8
$2.30	$2.42	$2.36	$2.49	$2.24	$2.36	$2.42	$2.49

■ **3-28** Industrial Suppliers, Inc. tiene registros del costo de procesamiento de cada pedido. Durante los últimos 5 años, este costo fue de $55.00, $58.00, $61.00, $65.00 y $66.00. ¿Cuál fue el crecimiento porcentual promedio de la empresa durante este lapso? Si esta tasa promedio se mantiene estable durante 3 años más, ¿cuánto le costará a la empresa procesar un pedido al final de ese periodo?

■ **3-29** Un sociólogo ha estado estudiando los cambios anuales en el número de convictos asignados al reclusorio más grande del estado. Sus datos están expresados en términos del aumento porcentual en el número de presos (un número negativo indica una disminución porcentual). Los datos más recientes recabados por el sociólogo son los siguientes:

1991	1992	1993	1994	1995	1996
−4%	5%	10%	3%	6%	−5%

a) Calcule el aumento porcentual promedio utilizando sólo los datos de 1992 a 1995.
b) Calcule el aumento porcentual promedio utilizando ahora los datos de los 6 años.
c) En 1990 se aprobó un nuevo código penal. Antes, la población del reclusorio crecía a una tasa de alrededor del 2% anual. ¿Cuál parece ser el efecto del nuevo reglamento?

Soluciones a los ejercicios de autoevaluación

EA 3-5 $\text{M.G.} = \sqrt[7]{1.11(1.09)(1.075)(1.08)(1.095)(1.108)(1.12)} = \sqrt[7]{1.908769992} = 1.09675$

El incremento promedio es 9.675% anual. La estimación de gastos por deudores morosos en 1997 es $(1.09675)^2 - 1 = 0.2029$, es decir, 20.29% más alto que en 1995.

EA 3-6 Los factores de crecimiento mensual son $275/300 = 0.9167$, $250/275 = 0.9091$, $240/250 = 0.9600$ y $225/240 = 0.9375$, de manera que

$$\text{M.G.} = \sqrt[4]{0.9167(0.9091)(0.9600)(0.9375)} = \sqrt[4]{0.7500} = 0.9306 = 1 - 0.0694$$

El precio ha disminuido a una tasa promedio del 6.94% mensual.

3.5 Una cuarta medida de tendencia central: la mediana

Definición de *mediana*

La *mediana* es una medida de tendencia central diferente a cualquiera de las que hemos tratado hasta ahora. La mediana es un solo valor del conjunto de datos que mide la observación central del conjunto. Esta sola observación es el elemento que está *más al centro* del conjunto de números. La mitad de los elementos están por arriba de este punto y la otra mitad está por debajo.

Cálculo de la mediana a partir de datos no agrupados

Localización de la mediana de datos no agrupados

Para hallar la mediana de un conjunto de datos, primero se organizan en orden descendente o ascendente. Si el conjunto de datos contiene un número *impar* de elementos, el de en medio en el arreglo es la mediana; si hay un número *par* de observaciones, la mediana es el promedio de los dos elementos de en medio. En lenguaje formal, la mediana es:

Mediana

$$\text{Mediana} = \left(\frac{n+1}{2}\right)\text{-ésimo término del arreglo de datos} \qquad [3\text{-}7]$$

(n: Número de elementos del arreglo)

Un número impar de elementos

Suponga que deseamos encontrar la mediana de siete elementos de un arreglo de datos. De acuerdo con la ecuación 3-7, la mediana es el cuarto término del arreglo $(7 + 1)/2 = 4$. Si aplicamos es-

Lo mediana no se ve distorsionada por valores extremos

to al ejemplo de los tiempos de los siete integrantes de un equipo de atletismo, descubriremos que el cuarto elemento del arreglo es 4.8 minutos. Ésta es la mediana del tiempo del equipo de atletismo. Observe que a diferencia de la media aritmética calculada, la mediana que calculamos en la tabla 3-l0 *no* se distorsiona por la presencia del último valor (9.0). Este valor pudo haber sido 15.0 o incluso 45.0, y la mediana ¡seguiría siendo la misma!

Un número par de elementos

Calculemos ahora la mediana de un arreglo con un número par de elementos. Considere los datos mostrados en la tabla 3-11 referentes al número de pacientes tratados diariamente en la sala de emergencias de un hospital. Los datos están organizados en orden descendente. La mediana de este conjunto de datos sería

$$\text{Mediana} = \left(\frac{n+1}{2}\right)\text{-ésimo término del arreglo de datos} \qquad [3\text{-}7]$$

$$= \frac{8+1}{2}$$

$$= 4.5\text{-ésimo término}$$

Como la mediana es el elemento número 4.5 del arreglo, necesitamos calcular el promedio de los elementos cuarto y quinto. El cuarto elemento de la tabla 3-11 es 43 y el quinto 35. El promedio de estos dos elementos es igual a $(43 + 35)/2 = 39$. Por consiguiente, 39 es la mediana del número de pacientes por día tratados en la sala de emergencias durante el periodo de 8 días.

Cálculo de la mediana a partir de datos agrupados

Búsqueda de la mediana de datos agrupados

A menudo, tenemos acceso a los datos hasta después de agruparlos en una distribución de frecuencias. Por ejemplo, no conocemos todas las observaciones que llevaron a la tabla 3-12, que contiene datos acerca de los 600 clientes bancarios considerados antes. En este caso, tenemos 10 intervalos de clase y un registro de las frecuencias con las que aparecen las observaciones en cada intervalo.

Localice la clase de la mediana

No obstante, podemos calcular la mediana del saldo de las cuentas de cheques de estos 600 clientes determinando cuál de los 10 intervalos de clase *contiene* la mediana. Para ello, debemos sumar las frecuencias que aparecen en la columna de frecuencias de la tabla 3-12 hasta que lleguemos al elemento número $(n + 1)/2$. Como tenemos 600 cuentas, el valor para $(n + 1)/2$ es 300.5 (el promedio de los números 300 y 301). El problema consiste en encontrar los intervalos de clase que contengan a los elementos número 300 y 301. La frecuencia acumulada para las dos primeras clases es sólo $78 + 123 = 201$. Pero cuando tomamos en cuenta al tercer intervalo de clase y sumamos 187 elementos a los 201 acumulados, tendremos un total de 388. En consecuencia, las observaciones número 300 y 301 deben estar en esta tercera clase (el intervalo de $100.00 a $149.99).

La *clase de la mediana* de este conjunto de datos contiene 187 observaciones. Si suponemos que estos 187 elementos empiezan en $100.00 y se encuentran *igualmente espaciados en todo el inter-*

Tabla 3-10

Tiempos para los integrantes de un equipo de atletismo

Elemento del arreglo de datos	1	2	3	4	5	6	7
Tiempo en minutos	4.2	4.3	4.7	4.8	5.0	5.1	9.0
				↑ Mediana			

Tabla 3-11

Pacientes tratados en la sala de urgencias durante 8 días consecutivos

Elemento del arreglo de datos	1	2	3	4	5	6	7	8
Número de pacientes	86	52	49	43	35	31	30	11
					↑ Mediana de 39			

Tabla 3-12	Clase en dólares	Frecuencia	
Saldos mensuales promedio de 600 clientes	0- 49.99	78	
	50.00- 99.99	123	
	100.00- 149.99	187	Clase de la mediana
	150.00-199.99	82	
	200.00-249.99	51	
	250.00-299.99	47	
	300.00-349.99	13	
	350.00-399.99	9	
	400.00-449.99	6	
	450.00-499.99	4	
		600	

valo de clase desde \$100.00 hasta \$149.99, entonces podemos interpolar y encontrar valores para los elementos 300 y 301. Primero determinamos que el elemento número 300 es la observación número 99 de la clase de la mediana:

$$300 - 201 \text{ [elementos de las primeras dos clases]} = 99$$

y que el elemento número 301 es la observación número 100 de la clase mediana:

$$301 - 201 = 100$$

Entonces podemos calcular el *ancho* de los 187 pasos iguales desde \$100.00 hasta \$149.99 de la siguiente manera:

Primer elemento de la siguiente clase — Primer elemento de la clase de la mediana

$$\frac{\$150.00 - \$100.00}{187} = \$0.267 \text{ de ancho}$$

Si existen 187 pasos de \$0.267 cada uno y necesitamos 98 pasos para llegar al elemento número 99, entonces éste es:

$$(\$0.267 \times 98) + \$100 = \$126.17$$

y el elemento número 100 está un paso más adelante:

$$\$126.17 + \$0.267 = \$126.44$$

Por tanto, podemos usar \$126.17 y \$126.44 como los valores de los elementos 300 y 301, respectivamente.

La mediana real de este conjunto de datos es el valor del elemento número 300.5, es decir, el promedio de las observaciones 300 y 301. Este promedio es:

$$\frac{\$126.17 + \$126.44}{2} = \$126.30$$

Esta cantidad (\$126.30) es la mediana de los saldos mensuales de las cuentas de cheques, estimada a partir de los datos agrupados de la tabla 3-12.

En resumen, podemos calcular la mediana de un conjunto de datos agrupados de la siguiente manera:

Pasos para encontrar la mediana de datos agrupados

1. Utilice la ecuación 3-7 para determinar qué observación de la distribución está más al centro (en este caso, el promedio de los elementos 300 y 301).
2. Sume las frecuencias de cada clase para encontrar la clase que contiene a ese elemento más al centro (la tercera clase, o \$100.00 − \$149.99).

3. Determine el número de elementos de la clase (187) y la localización de la clase que contiene a la mediana (la observación 300 fue el elemento número 99; la observación 301, el 100).
4. Determine el ancho de cada paso para pasar de una observación a otra en la clase mediana, dividiendo el intervalo de clase entre el número de elementos contenidos en la clase (ancho = \$0.267).
5. Determine el número de pasos que hay desde el límite inferior de la clase de la mediana hasta el elemento correspondiente a la mediana (98 pasos para el elemento número 99; 99 para el 100).
6. Calcule el valor estimado de la mediana multiplicando el número de pasos necesarios para llegar a la observación mediana por el ancho de cada paso y al producto súmele el valor del límite inferior de la clase mediana (\$100 + 98 × \$0.267 = \$126.17; \$126.17 + \$0.267 = \$126.44).
7. Si existe un número par de observaciones en la distribución, como en nuestro ejemplo, tome el promedio de los valores obtenidos para la mediana calculados en el paso número 6 (\$126.30).

Un método más sencillo

Para hacer más corto el procedimiento anterior, los especialistas en estadística utilizan una ecuación para determinar la mediana de un conjunto de datos agrupados. Para una muestra, la ecuación sería:

Mediana de la muestra para datos agrupados

$$\tilde{m} = \left(\frac{(n + 1)/2 - (F + 1)}{f_m}\right)w + L_m \qquad [3\text{-}8]$$

donde,

- $\tilde{m}$ = mediana de la muestra
- n = número total de elementos de la distribución
- F = suma de todas las frecuencias de clase *hasta*, pero *sin incluir*, la clase de la mediana
- f_m = frecuencia de la clase de la mediana
- w = ancho de intervalo de clase
- L_m = límite inferior del intervalo de clase de la mediana

Si utilizamos la ecuación 3-8 para calcular la mediana de nuestra muestra referente a los saldos de cuentas de cheques, entonces $n = 600$, $F = 201$, $f_m = 187$, $w = \$50$ y $L_m = \$100$.

$$\tilde{m} = \left(\frac{(n + 1)/2 - (F + 1)}{f_m}\right)w + L_m \qquad [3\text{-}8]$$

$$= \left(\frac{601/2 - 202}{187}\right)\$50 + \$100$$

$$= \left(\frac{98.5}{187}\right)\$50 + \$100$$

$$= (0.527)(\$50) + \$100$$

$$= \$126.35 \leftarrow \text{Mediana de la muestra estimada}$$

La pequeña diferencia entre este resultado y el que calculamos siguiendo el camino largo se debe al redondeo.

Ventajas y desventajas de la mediana

Ventajas de la mediana

La mediana tiene varias ventajas respecto a la media. La más importante, mostrada en el ejemplo del equipo de atletismo de la tabla 3-10, es que los valores extremos no afectan a la mediana de manera tan grave como a la media. La mediana es fácil de entender y se puede calcular a partir de cualquier tipo de datos, incluso a partir de datos agrupados con clases de extremo abierto como la distribución de frecuencias de la tabla 3-7, *a menos* que la mediana entre en una clase de extremo abierto.

Podemos encontrar la mediana incluso cuando nuestros datos son descripciones cualitativas como color o nitidez, en lugar de números. Suponga, por ejemplo, que tenemos tres tirajes de una prensa de imprenta. Los resultados deben clasificarse de acuerdo con la nitidez de la imagen. Podemos ordenar los resultados desde mejor hasta peor: extremadamente nítida, muy nítida, nítida, ligeramente borrosa y muy borrosa. La mediana de las cinco clasificaciones es la $(5 + 1)/2$, es decir la tercera (nítida).

Desventajas de la mediana

La mediana tiene también algunas desventajas. Ciertos procedimientos estadísticos que utilizan la mediana son más complejos que aquellos que utilizan la media. También, debido a que la mediana es una posición promedio, debemos ordenar los datos antes de llevar a cabo cualquier cálculo. Esto implica consumo de tiempo para cualquier conjunto de datos que contenga un gran número de elementos. Por consiguiente, si deseamos utilizar un estadístico de la muestra para estimar un parámetro de la población, la media es más fácil de usar que la mediana. En el capítulo 7 analizaremos el tema de la estimación con detalle.

SUGERENCIAS Y SUPOSICIONES

Hay buenas y malas noticias respecto al uso de la mediana. La buena noticia es que calcularla es bastante rápido y evita el efecto de valores muy grandes o muy pequeños. La mala noticia es que se sacrifica cierta exactitud al elegir un solo valor para representar una distribución. Para los valores 2, 4, 5, 40, 100, 213 y 347, la mediana es 40, que no tiene relación aparente con ninguno de los otros valores de la distribución. Advertencia: antes de hacer cálculos revise los datos con su propio sentido común. Si la distribución se ve poco usual, casi todo lo que calcule con esos datos tendrá defectos o limitaciones.

Ejercicios 3.5

Ejercicios de autoevaluación

EA 3-7 Swifty Markets compara los precios de artículos idénticos vendidos en sus tiendas de alimentos. Los precios siguientes, en dólares, corresponden a una libra de tocino, verificados la semana pasada.

$1.08 0.98 1.09 1.24 1.33 1.14 1.55 1.08 1.22 1.05

a) Calcule la mediana del precio por libra.
b) Calcule la media del precio por libra.
c) ¿Cuál es la mejor medida de tendencia central de estos datos?

EA 3-8 Para la siguiente distribución de frecuencias, determine:
a) La clase de la mediana.
b) El número de elemento que representa la mediana.
c) El ancho de los pasos iguales en la clase de la mediana.
d) El valor estimado de la mediana para estos datos.

Clase	Frecuencia	Clase	Frecuencia
100-149.5	12	300-349.5	72
150-199.5	14	350-399.5	63
200-249.5	27	400-449.5	36
250-299.5	58	450-499.5	18

Aplicaciones

■ **3-30** La empresa Meridian Trucking lleva un registro del kilometraje de todos sus vehículos. A continuación presentamos registros del kilometraje semanal:

810	450	756	789	210	657	589	488	876	689
1,450	560	469	890	987	559	788	943	447	775

a) Calcule la mediana del kilometraje que recorre un camión.
b) Calcule la media para el kilometraje de los 20 camiones.
c) Compare el resultado de los incisos a) y b) y explique cuál es la mejor medida de la tendencia central de los datos.

■ **3-31** El Consumer's Bureau de Carolina del Norte realizó una encuesta acerca de los proveedores de televisión por cable en el estado. Los siguientes datos se refieren al número de canales que ofrecen en el servicio básico:

32	28	31	15	25	14	12	29	22	28	29	32	33	24	26	8	35

a) Calcule la mediana del número de canales proporcionados.
b) Calcule el número medio de canales proporcionados.
c) ¿Qué valor es la mejor medida de tendencia central para estos datos?

■ **3-32** Para la siguiente distribución de frecuencias:
a) ¿Qué número representa la mediana?
b) ¿Qué clase contiene la mediana?
c) ¿Cuál es el ancho de los pasos iguales en la clase de la mediana?
d) ¿Cuál es el valor estimado de la mediana para estos datos?
e) Utilice la ecuación 3-8 para estimar la mediana de los datos. ¿Son cercanas entre sí sus dos estimaciones?

Clase	Frecuencia
10-19.5	8
20-29.5	15
30-39.5	23
40-49.5	37
50-59.5	46
60-69.5	52
70-79.5	84
80-89.5	97
90-99.5	16
100 o más	5

■ **3-33** Los siguientes datos representan el peso de los peces atrapados por el bote deportivo "El Fugitivo":

Clase	Frecuencia
0- 24.9	5
25- 49.9	13
50- 74.9	16
75- 99.9	8
100-124.9	6

a) Utilice la ecuación 3-8 para estimar la mediana del peso de los peces.
b) Utilice la ecuación 3-3 para calcular la media de estos datos.
c) Compare los incisos a) y b) y comente cuál es la mejor medida de tendencia central de los datos.

■ **3-34** El Departamento de Transporte de Chicago cree que el exceso de velocidad de los autobuses aumenta el costo de mantenimiento. Piensa que la mediana de los tiempos razonable para el recorrido del aeropuerto O'Hare al Centro John Hancock debería ser alrededor de 30 minutos. De la siguiente muestra de datos (en minutos) ¿puede usted ayudar al departamento a determinar si conducen los autobuses con exceso de velocidad? Si de los datos concluye que la velocidad fue excesiva, ¿qué explicación podrían darle los conductores de los autobuses?

17	32	21	22
29	19	29	34
33	22	28	33
52	29	43	39
44	34	30	41

■ **3-35** Mark Merritt, gerente de la Quality Upholstery Company, se encuentra investigando cantidad de material utilizado en los trabajos de tapicería de la empresa. La cantidad varía de un trabajo a otro, debido a los

diferentes estilos y tamaños de los muebles. Merrit reunió los datos siguientes (en yardas) de los trabajos hechos la semana anterior.

$5\frac{1}{4}$	$6\frac{1}{4}$	6	$7\frac{7}{8}$	$9\frac{1}{4}$	$9\frac{1}{2}$	$10\frac{1}{2}$
$5\frac{3}{8}$	6	$6\frac{1}{4}$	8	$9\frac{1}{2}$	$9\frac{7}{8}$	$10\frac{1}{4}$
$5\frac{1}{2}$	$5\frac{7}{8}$	$6\frac{1}{2}$	$8\frac{1}{4}$	$9\frac{3}{8}$	$10\frac{1}{4}$	$10\frac{1}{8}$
$5\frac{7}{8}$	$5\frac{3}{4}$	7	$8\frac{1}{2}$	$9\frac{1}{8}$	$10\frac{1}{2}$	$10\frac{1}{8}$
6	$5\frac{7}{8}$	$7\frac{1}{2}$	9	$9\frac{1}{4}$	$9\frac{7}{8}$	10

Si se tienen programados 150 trabajos para las siguientes 3 semanas, utilice la mediana para predecir cuántos metros de material se van a necesitar.

■ **3-36** Si la cantidad de reclamaciones por accidentes automovilísticos a una compañía de seguros muestra la siguiente distribución, determine la mediana utilizando el método descrito anteriormente. Verifique su resultado usando la ecuación 3-8.

Monto de reclamaciones ($)	Frecuencia	Monto de reclamaciones ($)	Frecuencia
menos que 250	52	750-999.99	1,776
250-499.99	337	1,000 o más	1,492
500-749.99	1,066		

■ **3-37** Un investigador obtuvo las respuestas siguientes a una de las preguntas incluidas en una encuesta de evaluación: totalmente en contra, en contra, ligeramente en contra, un poco de acuerdo, de acuerdo, altamente de acuerdo. De las seis respuestas, ¿cuál es la mediana?

Soluciones a los ejercicios de autoevaluación

EA 3-7 Primero se arreglan los precios en orden ascendente:

0.98 1.05 1.08 1.08 1.09 1.14 1.22 1.24 1.33 1.55

a) Mediana $= \dfrac{1.09 + 1.14}{2} = \$1.115$, el promedio de los datos 5 y 6

b) $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{11.76}{10} = \$1.176$

c) Debido a que los datos están ligeramente sesgados, la mediana puede ser un poco mejor que la media, pero en realidad no hay una diferencia notoria.

EA 3-8

Clase	Frecuencia	Frecuencia acumulada
100-149.5	12	12
150-199.5	14	26
200-249.5	27	53
250-299.5	58	111
300-349.5	72	183
350-399.5	63	246
400-449.5	36	282
450-499.5	18	300

a) Clase de la mediana = 300-349.5
b) Promedio de los datos 150 y 151
c) Ancho de paso = 50/72 = 0.6944
d) 300 + 38(0.6944) = 326.3872 (150)

$$300 + 39(0.6944) = \frac{327.0816}{653.4688} \quad (151)$$

$$\text{Mediana} = \frac{653.4688}{2} = 32.7344$$

3.6 Una medida final de tendencia central: la moda

Definición de *moda*

La moda es una medida de tendencia central diferente de la media, pero un tanto parecida a la mediana, pues en realidad no se calcula mediante algún proceso aritmético ordinario. La moda es *el valor que más se repite en el conjunto de datos*.

Riesgos al usar la moda de datos no agrupados

Como en todos los demás aspectos de la vida, el azar puede desempeñar un papel importante en la organización de datos. En ocasiones, el azar hace que un solo elemento no representativo se repita lo suficiente para ser el valor más frecuente del conjunto de datos. Es por esto que rara vez utilizamos la moda de un conjunto de datos no agrupados como una medida de tendencia central. La tabla 3-13, por ejemplo, presenta el número de viajes de entrega por día que hace una revolvedora de concreto. El valor modal es 15, ya que se presenta más a menudo que cualquier otro valor (tres veces). Una moda de 15 implica que la actividad de la planta es mayor que 6.7 (6.7 es el resultado al calcular la media). La moda nos dice que 15 es el número más frecuente de viajes, pero no nos indica que la mayor cantidad de viajes está por debajo de 10.

Búsqueda de la clase modal de datos agrupados

Agrupemos ahora estos datos en una distribución de frecuencias, como en la tabla 3-14. Si seleccionamos la clase que tiene el mayor número de observaciones, a la cual podemos llamar *clase modal*, elegimos 4-7 viajes. Esta clase es más representativa de la actividad de la revolvedora que la moda de 15 viajes diarios. Por esto, siempre que utilizamos la moda como una medida de tendencia central de un conjunto de datos, debemos calcular la moda de datos agrupados.

Cálculo de la moda de datos agrupados

Cuando los datos ya se encuentran agrupados en una distribución de frecuencias, podemos suponer que la moda está localizada en la clase que contiene el mayor número de elementos, es decir, en la clase que tiene la mayor frecuencia. Para determinar un solo valor para la moda a partir de esta clase modal, utilizamos la ecuación 3-9:

Moda

$$Mo = L_{Mo} + \left(\frac{d_1}{d_1 + d_2}\right)w \qquad [3\text{-}9]$$

donde,

- L_{M_O} = límite inferior de la clase modal
- d_1 = frecuencia de la clase modal menos la frecuencia de la clase que se encuentra *inmediatamente menor que ella*

Tabla 3-13
Viajes de entrega por día en un periodo de 20 días

Viajes organizados en orden ascendente					
0	2	5	7	15	} ← Moda
0	2	5	7	15	
1	4	6	8	15	
1	4	6	12	19	

Tabla 3-14
Distribución de frecuencias de los viajes de entrega

Clase de número de entregas	0-3	4-7	8-11	12 o más
Frecuencia	6	8	1	5
		↑ Clase modal		

- $d_2 =$ frecuencia de la clase modal menos la frecuencia de la clase *inmediatamente mayor que ella*
- $w =$ ancho del intervalo de la clase modal

Si utilizamos la ecuación 3-9 para calcular la moda del saldo de las cuentas de cheques de nuestro ejemplo (vea la tabla 3-12), entonces $L_{Mo} = \$100$, $d_1 = 187 - 123 = 64$, $d_2 = 187 - 82 = 105$ y $w = \$50$.

$$Mo = L_{Mo} + \left(\frac{d_1}{d_1 + d_2}\right)w \qquad [3\text{-}9]$$

$$= \$100 + \frac{64}{64 + 105}\$50$$

$$= \$100 + (0.38)(\$50)$$

$$= \$100 + \$19$$

$$= \$119.00 \leftarrow \text{Moda}$$

El resultado obtenido, \$119, es la estimación de la moda.

Distribuciones multimodales

Distribuciones bimodales

¿Qué sucede cuando tenemos dos valores diferentes y *cada uno* parece ser el mayor número de veces que aparece un valor en un conjunto de datos? En la tabla 3-15 se muestran los errores de facturación en un periodo de 20 días cometidos en las oficinas administrativas de un hospital. Observe que tanto 1 como 4 parecen ser el mayor número de errores del conjunto de datos. Ambos aparecen tres veces. Esta distribución, entonces, tiene dos modas y se le conoce como *distribución bimodal*.

En la figura 3-6, se graficaron los datos de la tabla 3-15. Observe que hay dos puntos que son los más altos de la gráfica. Se presentan con los valores correspondientes a 1 y 4 errores de facturación. La distribución de la figura 3-7 se conoce también como bimodal, aunque en este caso los dos valores más altos no sean iguales. Es claro que estos puntos son mayores que los valores más cercanos de la frecuencia observada.

Tabla 3-15

Errores de facturación por día en un periodo de 20 días

Errores organizados en orden ascendente			
0	2	6	9
0	4	6	9
1 ← Moda	4 ← Moda	7	10
1 ← Moda	4	8	12
1	5	8	12

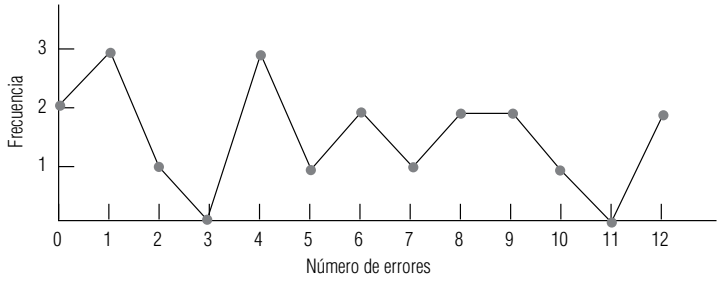


FIGURA 3-6

Datos de la tabla 3-15 que muestran una distribución bimodal

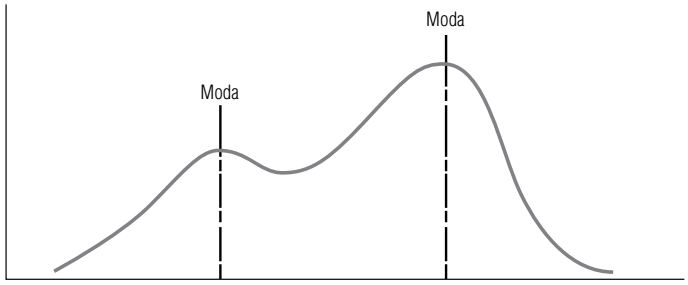


FIGURA 3-7

Distribución bimodal con dos modas distintas

Ventajas y desventajas de la moda

Ventajas de la moda

La moda, igual que la mediana, se puede utilizar como una posición central para datos tanto cualitativos como cuantitativos. Si una prensa estampa cinco impresiones que podemos clasificar como "muy nítida", "nítida", "nítida", "nítida" y "borrosa", entonces el valor modal es "nítida". De manera análoga, podemos hablar de estilos modales cuando, por ejemplo, los clientes de una mueblería prefieren muebles tipo "colonial" sobre cualquier otro estilo.

También, al igual que la mediana, **los valores extremos no afectan indebidamente a la moda**. Aun cuando los valores extremos sean muy altos o muy bajos, escogemos el valor más frecuente del conjunto de datos como el valor modal. Podemos utilizar la moda sin importar qué tan grandes o qué tan pequeños sean los valores del conjunto de datos e independientemente de cuál sea su dispersión.

Una tercera ventaja de la moda es que la podemos utilizar aun cuando una o más clases sean de extremo abierto. Note, por ejemplo, que la tabla 3-14 contiene la clase de extremo abierto "12 viajes o más".

Desventajas de la moda

A pesar de estas ventajas, la moda no se utiliza tan a menudo como medida de tendencia central, como se usan la media y la mediana. Muchas veces, no existe un valor modal debido a que el conjunto de datos no contiene valores que se presenten más de una vez. En otras ocasiones, cada valor es la moda, pues cada uno de ellos se presenta el mismo número de veces. Resulta claro que la moda es una medida inútil en tales casos. Otra desventaja consiste en que cuando los conjuntos de datos contienen dos, tres o más modas, es difícil interpretarlos y compararlos.

Comparación de la media, la mediana y la moda

La media, la mediana y la moda son idénticas en una distribución simétrica

Cuando trabajamos problemas de estadística, debemos decidir si vamos a utilizar la media, la mediana o la moda como medidas de tendencia central. Las distribuciones simétricas que sólo contienen una moda siempre tienen el mismo valor para la media, la mediana y la moda. En esos casos, no es necesario escoger la medida de tendencia central, pues ya está hecha la selección.

En una distribución con sesgo positivo (es decir, sesgada a la derecha), como la gráfica (a) de la figura 3-8, la moda todavía se encuentra en el punto más alto de la distribución, la mediana está a la derecha de la moda y la media se encuentra todavía más a la derecha de la moda y la mediana.

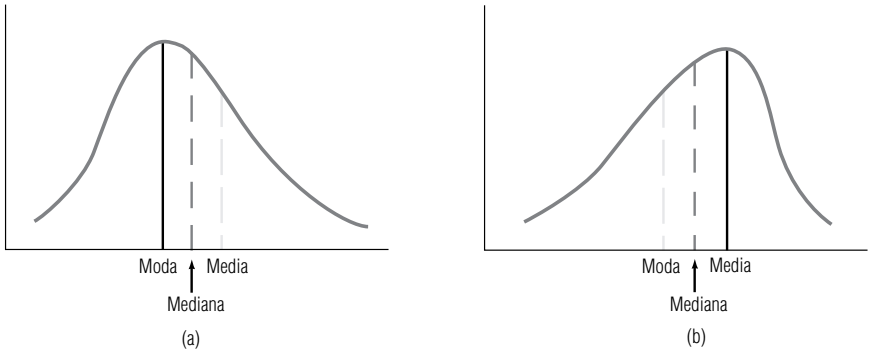


FIGURA 3-8

Distribuciones con sesgo (a) positivo y (b) negativo que muestran las posiciones de la media, la mediana y la moda

En una distribución con sesgo negativo (es decir, sesgada a la izquierda), como en la gráfica (b) de la figura 3-8, la moda sigue siendo el punto más alto de la distribución, la mediana está a la izquierda y la media se encuentra todavía más a la izquierda de la moda y la mediana.

La mediana puede ser la mejor medida de posición en distribuciones sesgadas

Cuando la población está sesgada negativa o positivamente, la mediana suele ser la mejor medida de posición, debido a que siempre está entre la moda y la media. La frecuencia de ocurrencia de un solo valor no influye mucho en la mediana como es el caso de la moda, ni la distorsionan los valores extremos como la media.

En cualquier otro caso, no existen guías universales para la aplicación de la media, la mediana o la moda como medidas de tendencia central para diferentes poblaciones. Cada caso deberá considerarse de manera independiente, de acuerdo con las líneas generales que se analizaron.

SUGERENCIAS Y SUPOSICIONES

Sugerencia: al intentar decidir los usos de la media, la mediana y la moda, debe pensarse en las situaciones prácticas en las que cada una de ellas tiene más sentido. Si se obtiene el promedio de un pequeño grupo de salarios en una fábrica bastante cercanos entre sí, la media aritmética es muy exacta y se calcula rápidamente. Si existen 500 casas nuevas en un desarrollo urbano, todas con valores que no difieren en más de $10,000, entonces la mediana es mucho más rápida y también bastante exacta. Al manejar los efectos acumulados de la inflación o el interés, se requiere la media geométrica si se desea exactitud. Un ejemplo de sentido común: aunque es cierto que la familia promedio tiene 1.65 hijos, los diseñadores de automóviles tomarán mejores decisiones si usan el valor modal de 2.0 niños.

Ejercicios 3.6

Ejercicios de autoevaluación

EA 3-9 Las siguientes son las edades en años de los automóviles en los que trabajó Village Autohaus la semana pasada:

5 6 3 6 11 7 9 10 2 4 10 6 2 1 5

a) Calcule la moda para este conjunto de datos.
b) Calcule la media para este conjunto de datos.
c) Compare los incisos a) y b) y comente cuál es la mejor medida de tendencia central de estos datos.

EA 3-10 Las edades de una muestra de estudiantes que asisten a Sandhills Community College este semestre son:

19	17	15	20	23	41	33	21	18	20
18	33	32	29	24	19	18	20	17	22
55	19	22	25	28	30	44	19	20	39

a) Construya una distribución de frecuencias con intervalos 15-19, 20-24, 25-29, 30-34 y 35 o más.
b) Estime el valor de la moda mediante la ecuación 3-9.
c) Ahora calcule la media de los datos sin procesar.
d) Compare sus repuestas a los incisos b) y c) y comente cuál de las dos medidas de tendencia central es más adecuada para estos datos y por qué.

Aplicaciones

■ **3-38** Un bibliotecario encuestó a 20 personas al salir de la biblioteca y les preguntó cuántos libros habían sacado. Las respuestas fueron las siguientes:

1 0 2 2 3 4 2 1 2 0 2 2 3 1 0 7 3 5 4 2

a) Calcule la moda de este conjunto de datos.
b) Calcule la media para este conjunto de datos.
c) Grafique los datos de la frecuencia contra el número de libros sacados. ¿Es la media o la moda una mejor medida de tendencia central para estos datos?

■ **3-39** La edad de los residentes de Twin Lakes Retirement Village tiene la siguiente distribución de frecuencias:

Clase	Frecuencia
47-51.9	4
52-56.9	9
57-61.9	13
62-66.9	42
67-71.9	39
72-76.9	20
77-81.9	9

Estime el valor modal de la distribución utilizando la ecuación 3-9.

■ **3-40** ¿Cuáles son los valores modales para las siguientes distribuciones?

(a) Color de cabello	Negro	Castaño	Pelirrojo	Rubio			
Frecuencia	11	24	6	18			
(b) Tipo de sangre	AB	O	A	B			
Frecuencia	4	12	35	16			
(c) Día de nacimiento	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado	Domingo
Frecuencia	22	10	32	17	13	32	14

■ **3-41** Los siguientes datos se refieren al número de departamentos en 27 complejos en la ciudad de Cary, Carolina del Norte.

91	79	66	98	127	139	154	147	192
88	97	92	87	142	127	184	145	162
95	89	86	98	145	129	149	158	241

a) Construya una distribución de frecuencias usando los intervalos 66-87, 88-109, . . . , 220-241.
b) Estime el valor de la moda con la ecuación 3-9.
c) Calcule la media de los datos sin procesar.
d) Compare sus respuestas a los incisos b) y c) y comente cuál de las dos es mejor medida de tendencia central para estos datos y por qué.

■ **3-42** Estime la moda de la distribución dada en el ejercicio 3-36.

■ **3-43** El número de sistemas de calentamiento solar disponibles al público es bastante grande y su capacidad de almacenamiento de calor, diversa. A continuación presentamos una distribución de la capacidad de almacenamiento de calor (en días) de 28 sistemas que fueron probados recientemente por University Laboratories, Inc.:

Días	Frecuencia
0-0.99	2
1-1.99	4
2-2.99	6
3-3.99	7
4-4.99	5
5-5.99	3
6-6.99	1

En los laboratorios, se sabe que el informe sobre las pruebas circulará ampliamente y se usará como base para una legislación sobre los impuestos a las concesiones de los sistemas. En consecuencia, se desea que las medidas utilizadas reflejen los datos tanto como sea posible.

a) Calcule la media del conjunto de datos.
b) Calcule la moda del conjunto de datos.
c) Calcule la mediana del conjunto de datos.
d) Seleccione la respuesta entre los resultados de los incisos a), b) y c) que mejor refleje la tendencia central de los datos y justifique su elección.

■ **3-44** Ed Grant es director de la Oficina de Becas Estudiantiles del Wilderness College. Con datos disponibles acerca de los ingresos obtenidos en el verano por todos los estudiantes que han solicitado ayuda económica a la oficina, desarrolló la distribución de frecuencias siguiente:

Ingresos en el verano	Número de estudiantes
$ 0- 499	231
500- 999	304
1,000-1,499	400
1,500-1,999	296
2,000-2,499	123
2,500-2,999	68
3,000 o más	23

a) Encuentre la clase modal del conjunto de datos.
b) Utilice la ecuación 3-9 para encontrar la moda de los datos que utilizó Ed.
c) Si las becas a los estudiantes están restringidas a aquellos cuyos ingresos en el verano fueron por lo menos 10% menores que la ganancia modal, ¿cuántos solicitantes obtienen la beca?

Soluciones a los ejercicios de autoevaluación

EA 3-9 a) Moda = 6

b) $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{87}{15} = 5.8$

c) Como la frecuencia modal es sólo 3 y los datos son razonablemente simétricos, la media es mejor medida de tendencia central.

EA 3-10 a)

Clase	15-19	20-24	25-29	30-34	≥ 35
Frecuencia	10	9	3	4	4

b) $Mo = L_{M_o} + \dfrac{d_1}{d_1 + d_2} w = 15 + \left(\dfrac{10}{10 + 1}\right)5 = 19.55$

c) $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{760}{30} = 25.33$

d) Debido a que esta distribución está muy sesgada, la moda es una mejor medida de tendencia central.

3.7 Dispersión: por qué es importante

Necesidad de medir la dispersión o lo variabilidad

Al inicio de este capítulo, en la figura 3-2, mostramos dos conjuntos de datos con la misma posición central, pero uno con mayor dispersión que el otro. Esto sucede también con las tres distribuciones de la figura 3-9. La media de las tres curvas es la misma, pero la curva A tiene menor separación (o *variabilidad*) que la curva B, y ésta tiene menor variabilidad que la C. Si medimos sólo la media de estas tres distribuciones, estaremos pasando por alto una diferencia importante que existe entre las tres curvas. Al igual que sucede con cualquier conjunto de datos, la media, la mediana y la moda sólo nos revelan una parte de la información que debemos conocer acerca de las características de los

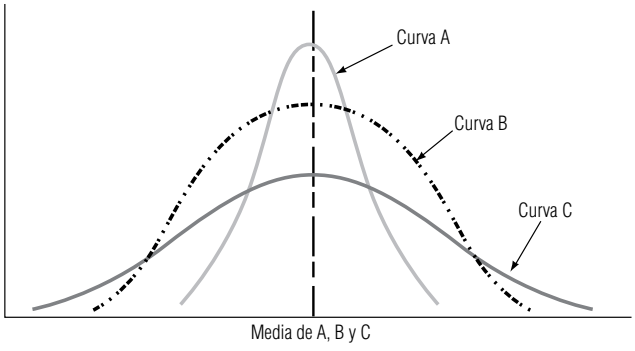


FIGURA 3-9

Tres curvas con la misma media pero diferente variabilidad

datos. Para aumentar nuestro entendimiento del patrón de los datos, debemos medir también su *dispersión*, separación o variabilidad.

Usos de las medidas de dispersión

¿Por qué es tan importante entender y medir la dispersión de la distribución? **Primero**, nos proporciona información adicional que nos permite juzgar la confiabilidad de nuestra medida de tendencia central. Si los datos se encuentran muy dispersos, como los que representa la curva C de la figura 3-9, la posición central es menos representativa de los datos, como un todo, que cuando éstos se agrupan más cerca alrededor de la media, como en la curva A de la misma figura. **Segundo**, ya que existen problemas característicos para datos muy dispersos, debemos ser capaces de reconocer esa dispersión amplia para poder abordar esos problemas. **Tercero**, quizá se desee comparar las dispersiones de diferentes muestras. Si no se desea tener una amplia dispersión de valores con respecto del centro de distribución, o esto presenta riesgos inaceptables, necesitamos poder reconocerla y evitar elegir distribuciones que tengan las dispersiones más grandes.

Usos financiero y en control de la calidad

Los analistas financieros están preocupados por la dispersión de las ganancias de una empresa. Las ganancias ampliamente dispersas —que van desde extremadamente altas a extremadamente bajas e incluso a niveles negativos— son indicativas de un riesgo mayor para los accionistas y para los acreedores que las ganancias que permanecen relativamente estables. De manera similar, los expertos en el control de la calidad analizan la dispersión de los niveles de calidad de un producto. Una medicina cuya pureza promedio es buena, pero que oscila desde muy pura hasta muy impura puede ser peligrosa para la vida humana.

SUGERENCIAS Y SUPOSICIONES

Los fabricantes de asientos para aviones hacen una suposición de la forma del viajero promedio. En algunas secciones de clase turista es común encontrar anchos de asientos de sólo 19″. Para alguien que pesa 250 libras (cerca de 113 kg) y usa talla 44, sentarse en un asiento de 19″ es como ponerse un zapato apretado. En el fútbol americano, por otro lado, ignorar la dispersión de los datos puede causar problemas graves. Un equipo que en promedio recorre 3.6 yardas por jugada, en teoría, debe ganar todos los juegos porque 3.6×4 jugadas es más que las 10 yardas necesarias para conservar el balón. Sin embargo, un poco de mala suerte y una pérdida ocasional de 20 yardas, afectan al *invencible* promedio teórico de 3.6 yardas. Advertencia: no invierta mucho en promedios a menos que sepa que la dispersión es pequeña. Un reclutador de la Fuerza Aérea de Estados Unidos que busca capacitar pilotos que en promedio midan 6 pies (1.82 m), quedaría despedido si se presenta con un aspirante de 4 pies (1.22 m) de estatura y otro de 8 pies (2.43 m). En la cláusula "razón de despido" de su expediente personal deberá decir "ignoró la dispersión".

Ejercicios 3.7

Conceptos básicos

■ **3-45** ¿Para cuál de las siguientes distribuciones la media es más representativa de los datos como un todo? ¿Por qué?

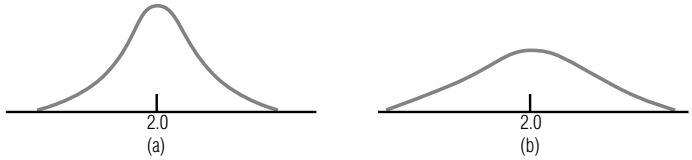
■ **3-46** ¿Cuál de las siguientes no es una razón válida para medir la dispersión de una distribución?

a) Indica la confiabilidad del estadístico empleado para medir la tendencia central.
b) Permite comparar varias muestras con promedios similares.
c) Utiliza más datos para describir una distribución.
d) Atrae la atención respecto a problemas asociados con distribuciones que tienen una variabilidad muy grande o muy pequeña.

Aplicaciones

■ **3-47** Para medir el éxito escolar, los educadores necesitan probar los niveles de conocimientos y habilidades de los estudiantes. Tomar en cuenta las diferencias individuales de cada uno de ellos, permite a los profesores planear mejor el programa académico. Las curvas que se muestran a continuación representan las distribuciones basadas en resultados anteriores de dos pruebas distintas. ¿Cuál de ellas seleccionaría usted como mejor opción para los propósitos de los profesores?

■ **3-48** Una empresa que usa dos métodos diferentes para enviar pedidos a sus clientes encontró las siguientes distribuciones del tiempo de entrega para los dos métodos, según los registros históricos. Con la evidencia disponible, ¿qué método de envío recomendaría?



■ **3-49** De las tres curvas de la figura 3-9, escoja la que sirva mejor para describir la distribución de las edades de los grupos siguientes: miembros del Congreso; miembros recientemente electos de la Cámara de Diputados; y presidentes de las diferentes comisiones de la misma cámara. Al hacer su elección, no tome en cuenta la media de las curvas de la figura 3-9 y considere sólo la variabilidad de la distribución. Establezca brevemente las razones que lo llevaron a elegir esas distribuciones.

■ **3-50** ¿De qué manera cree que debe aplicarse el concepto de variabilidad a una investigación que realiza la Secretaría de Comercio (SC) con el propósito de determinar la posibilidad de que un grupo de fabricantes fije los precios de los productos?

■ **3-51** Escoja cuál de las tres curvas que se muestran en la figura 3-9 describe mejor la distribución de las siguientes características de diferentes grupos. Haga sus elecciones con base sólo en la variabilidad de las distribuciones. Explique brevemente la razón de cada elección.

a) El número de puntos obtenidos por cada jugador de una liga profesional de básquetbol durante la temporada de 80 juegos.
b) El salario de cada una de las 100 personas que trabajan en empleos aproximadamente iguales en el gobierno federal.
c) El promedio de calificaciones de cada uno de los 15,000 estudiantes de una universidad estatal.
d) El salario de cada una de las 100 personas que trabajan en empleos aproximadamente iguales en una empresa privada.
e) El promedio de calificaciones de cada estudiante de una universidad estatal que ha sido aceptado en el posgrado.
f) El porcentaje de tiros a la canasta lanzados por cada jugador de una liga profesional de básquetbol durante la temporada de 80 juegos.

3.8 Rangos: medidas de dispersión útiles

Tres medidas de distancia

La dispersión puede medirse en términos de la diferencia entre dos valores seleccionados del conjunto de datos. En esta sección, estudiaremos tres de las llamadas *medidas de distancia*: el rango, el rango interfractil y el rango intercuartil.

Tabla 3-16

Pagos anuales hechos por Blue Cross-Blue Shield (miles)

Cumberland	863	903	957	1,041	1,138	1,204
	1,354	1,624	1,698	1,745	1,802	1,883
Valley falls	490	540	560	570	590	600
	610	620	630	660	670	690

Rango

Definición y cálculo del rango

El *rango* es la diferencia entre el más alto y el más pequeño de los valores observados. En forma de ecuación, podemos decir

Rango

$$\text{Rango} = \begin{matrix}\text{valor de la observación} \\ \text{más grande}\end{matrix} - \begin{matrix}\text{valor de la observación} \\ \text{más pequeña}\end{matrix} \qquad [3\text{-}10]$$

Utilizando esta ecuación, podemos comparar los rangos de los pagos anuales que hace la asociación Blue Cross-Blue Shield a dos hospitales presentados en la tabla 3-16.

El rango de los pagos anuales a Cumberland es \$1,883,000 − \$863.000 = \$1,020,000. Para el hospital Valley Falls, el rango es \$690,000 − \$490,000 = \$200,000.

Características del rango

Es fácil entender y encontrar el rango, pero su utilidad como medida de dispersión es limitada. El rango sólo toma en cuenta los valores más alto y más bajo de una distribución y ninguna otra observación del conjunto de datos. Como resultado, ignora la naturaleza de la variación entre todas las demás observaciones, y tiene una gran influencia de los valores extremos. Debido a que sólo mide dos valores, el rango tiene muchas posibilidades de cambiar drásticamente de una muestra a la siguiente en una población dada, aunque los valores que caen entre el más alto y el más bajo sean bastante parecidos. Recuerde también que las distribuciones de extremo abierto no tienen rango, pues no existe un valor "más alto" o "más bajo" en la clase de extremo abierto.

Rango interfractil

Fractiles

En una distribución de frecuencias, una fracción o proporción dada de los datos cae en un *fractil* o abajo de éste. La mediana, por ejemplo, es el fractil 0.5, porque la mitad del conjunto de datos es menor o igual que este valor. Se dará cuenta que los fractiles son parecidos a los porcentajes. En una

Significado del rango interfractil

distribución cualquiera, 25% de los datos están en el fractil 0.25 o abajo de éste; igualmente, 25% de los datos cae en el vigésimo quinto percentil o es menor que éste. El *rango interfractil* es una medida de la dispersión entre dos fractiles de una distribución de frecuencias, es decir, la diferencia entre los valores de los dos fractiles.

Cálculo del rango interfractil

Suponga que deseamos encontrar el rango interfractil entre el primero y segundo tercios de los donativos recibidos por Cumberland de la organización Blue Cross-Blue Shield. Empezamos por dividir las observaciones en tercios, como en la tabla 3-17. Cada tercio contiene cuatro observaciones (.33% del total de 12 elementos). Entonces, 33.33% de los elementos está en \$1,041,000 o abajo de

Tabla 3-17

Pagos anuales de la Blue Cross-Blue Shield al Hospital Cumberland (miles)

Primer tercio	Segundo tercio	Último tercio
863	1,138	1,698
903	1,204	1,745
957	1,354	1,802
1,041 ← $^1/_3$ fractil	1,624 ← $^2/_3$ fractil	1,883

este valor, y 66.66% es menor o igual que $1,624,000. Ahora podemos calcular el rango interfractil entre los fractiles .33 y .66 restando $1,0141,000 del valor $1,624,000. Esta diferencia de $583,000 es la dispersión entre el valor más alto del primer tercio de los pagos y el valor más alto del segundo tercio.

Fractiles especiales: deciles, cuartiles y percentiles

Los fractiles tienen nombres especiales, dependiendo del número de partes iguales en que dividen a los datos. Los fractiles que los dividen en 10 partes iguales se llaman *deciles*. Los *cuartiles* dividen los datos en cuatro partes iguales. Los *percentiles* dividen al conjunto de datos en 100 partes iguales.

Rango intercuartil

Cálculo del rango intercuartil

El rango intercuartil mide aproximadamente qué tan lejos de la mediana debemos ir en cualquiera de las dos direcciones antes de recorrer una mitad de los valores del conjunto de datos. Para calcular este rango, dividimos nuestros datos en cuatro partes, cada una de las cuales contiene 25% de los elementos de la distribución. Los *cuartiles* son, entonces, los valores más altos de cada una de estas cuatro partes, y el *rango intercuartil* es la diferencia entre los valores del primero y tercer cuartiles:

Rango intercuartil

$$\text{Rango intercuartil} = Q_3 - Q_1 \qquad [3\text{-}11]$$

En la figura 3-10 se ilustra el concepto de rango intercuartil. Observe que los anchos de los cuatro cuartiles *no* necesariamente son los mismos.

En la figura 3-11, otra presentación de cuartiles donde éstos dividen el área bajo la distribución en cuatro partes iguales, cada una contiene 25% del área.

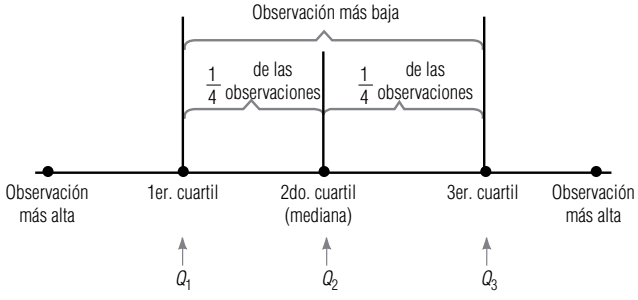


FIGURA 3-10

Rango intercuartil

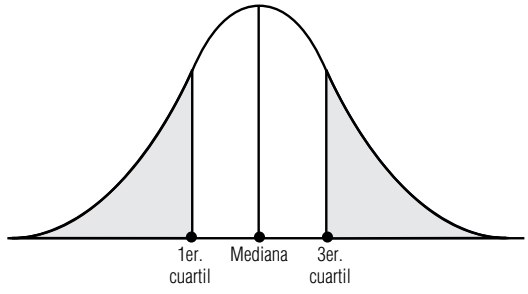


FIGURA 3-11

Cuartiles

SUGERENCIAS Y SUPOSICIONES

Fractil es un término que usan más los estadísticos que el resto de las personas, más familiarizadas con 100 fractiles o percentiles, en especial cuando se trata del percentil de la calificación en los exámenes de aptitud académica o de admisión a las universidades. Cuando se obtiene una letra que indica que el percentil de la calificación es 35, se sabe que 35% de quienes presentaron el examen lo hicieron peor que uno. Es más fácil comprender el significado del intervalo en especial cuando el profesor publica las calificaciones más altas y más bajas del siguiente examen de estadística. Sugerencia: todos estos términos ayudan a manejar la dispersión de los datos. Si todos los valores se ven parecidos, entonces el tiempo dedicado a calcular los valores de dispersión quizá no valga mucho. Si los datos se dispersan mucho, será riesgoso apostar al promedio sin considerar la dispersión.

Ejercicios 3.8

Ejercicios de autoevaluación

EA 3.11 Se presentan las calificaciones de un examen de historia. Encuentre el percentil 80.

95	81	59	68	100	92	75	67	85	79
71	88	100	94	87	65	93	72	83	91

EA 3.12 La compañía Casual Life Insurance estudia la compra de una nueva flota de autos. El director del Departamento de Finanzas, Tom Dawkins, obtuvo una muestra de 40 empleados para determinar el número de millas que cada uno maneja en un año. Los resultados del estudio son los siguientes. Calcule el rango y el rango intercuartil.

3,600	4,200	4,700	4,900	5,300	5,700	6,700	7,300
7,700	8,100	8,300	8,400	8,700	8,700	8,900	9,300
9,500	9,500	9,700	10,000	10,300	10,500	10,700	10,800
11,000	11,300	11,300	11,800	12,100	12,700	12,900	13,100
13,500	13,800	14,600	14,900	16,300	17,200	18,500	20,300

Conceptos básicos

■ **3-52** Para los siguientes datos, calcule el rango intercuartil.

99	75	84	61	33	45	66	97	69	55
72	91	74	93	54	76	52	91	77	68

■ **3-53** Para la muestra siguiente, calcule

a) el rango,
b) el rango interfractil entre los percentiles 20 y 80,
c) el rango intercuartil.

2,549	3,897	3,661	2,697	2,200	3,812	2,228	3,891	2,668	2,268
3,692	2,145	2,653	3,249	2,841	3,469	3,268	2,598	3,842	3,362

Aplicaciones

■ **3-54** Se dan las lecturas de temperaturas altas durante junio de 1995 en Phoenix, Arizona. Encuentre el percentil 70

84	86	78	69	94	95	94	98	89	87	88	89	92	99	102
94	92	96	89	88	87	88	84	82	88	94	97	99	102	105

■ **3-55** Los siguientes son los ingresos totales por viajes (en dólares) recolectados un martes por 20 taxis que pertenecen a City Transit, Ltd.

147	95	93	127	143	101	123	83	135	129
185	92	115	126	157	93	133	51	125	132

Calcule el rango de estos datos y comente si piensa que es una medida de dispersión útil.

■ **3-56** La empresa Redi-Mix Incoporated elaboró el siguiente registro del tiempo (redondeado a centésimos de minuto) que esperan sus camiones para la descarga en la obra. Calcule el rango y el rango intercuartil.

0.10	0.45	0.50	0.32	0.89	1.20	0.53	0.67	0.58	0.48
0.23	0.77	0.12	0.66	0.59	0.95	1.10	0.83	0.69	0.51

■ **3-57** La Warlington Appliances ha desarrollado una nueva combinación de mezcladora-vasija. Mediante una demostración de mercadotecnia y una investigación de precios, se determina que la mayoría de las perso-

nas muestreadas estaría dispuesta a pagar aproximadamente $60 por ella, con un rango intercuartil, sorpresivamente pequeño de $14. En un intento por obtener los mismos resultados, la demostración y la investigación correspondiente se repitieron. El departamento de mercadotecnia espera encontrar un rango intercuartil más pequeño. Los siguientes son los datos que se obtuvieron. ¿La esperanza del departamento se hizo realidad?

52	35	48	46	43	40	61	49	57	58	65	46
72	69	38	37	55	52	50	31	41	60	45	41
55	38	51	49	46	43	64	52	60	61	68	49
69	66	35	34	52	49	47	28	38	57	42	38

■ **3-58** MacroSwift ha decidido desarrollar un nuevo programa de software diseñado para directores ejecutivos y otros altos niveles. La compañía no desea desarrollar un programa que requiera demasiado espacio en el disco duro, por lo que sondearon a 36 ejecutivos para determinar la cantidad de espacio disponible en sus computadoras. Los resultados en megabytes son los siguientes:

6.3	6.7	7.9	8.4	9.7	10.6	12.4	19.4	29.1	42.6
59.8	97.6	100.4	120.6	135.5	148.6	178.6	200.1	229.6	284.6
305.6	315.6	325.9	347.5	358.6	397.8	405.6	415.9	427.8	428.6
439.5	440.9	472.3	475.9	477.2	502.6				

Calcule el rango y el rango intercuartil.

■ **3-59** El Departamento de Carreteras de Nuevo México tiene la tarea de mantener en buen estado todos los caminos estatales. Una medida de la condición de una carretera es el número de grietas que presenta por cada 30 metros de recorrido. A partir de la muestra anual que hace el departamento, se obtuvieron los siguientes datos:

4	7	8	9	9	10	11	12	12	13
13	13	13	14	14	14	15	15	16	16
16	16	16	17	17	17	18	18	19	19

Calcule los rangos interfractiles entre los percentiles 20, 40, 60 y 80.

■ **3-60** Ted Nichol es un analista estadístico que trabaja para los altos mandos administrativos de Research Incorporated. Ayudó a diseñar el lema publicitario de la compañía: "Si no puede encontrar la respuesta, entonces ¡INVESTÍGUELA!" Ted acaba de recibir algunos datos que le preocupan, el volumen mensual en dólares de los contratos de investigaciones que la compañía firmó durante el año anterior. Idealmente, estas cantidades mensuales deberían ser bastante estables, debido a que una fluctuación demasiado grande en la cantidad de trabajo a realizar puede tener como resultado una cantidad extraordinaria de contrataciones y despidos de empleados. Los datos de Ted (en miles de dólares) son los siguientes:

253	104	633	57	500	201
43	380	467	162	220	302

Calcule lo siguiente:

a) El rango interfractil entre los deciles 2 y 8.
b) La mediana, Q_1 y Q_3.
c) El rango intercuartil.

Soluciones a los ejercicios de autoevaluación

EA 3.11 Primero, se ordenan los datos en orden ascendente.

59	65	67	68	71	72	75	79	81	83
85	87	88	91	92	93	94	95	100	100

El dato 16 (es decir 93) es el percentil 80.

EA 3.12 Rango = 20,300 − 3,600 = 16,700 millas
Rango intercuartil = $Q_3 - Q_1$ = 12,700 − 8,100 = 4,600 millas.

3.9 Dispersión: medidas de desviación promedio

Dos medidas de desviación promedio

Las descripciones más completas de la dispersión son aquellas que manejan la desviación promedio respecto a alguna medida de tendencia central. Dos de estas medidas son importantes para nuestro estudio de la estadística: la *varianza* y la *desviación estándar*. Ambas medidas nos dan una distancia promedio de cualquier observación del conjunto de datos respecto a la media de la distribución.

Varianza de población

Varianza

Cada población tiene una varianza, su símbolo es σ^2 (sigma cuadrada). Para calcular la varianza de una población, la suma de los cuadrados de las distancias entre la media y cada elemento de la población se divide entre el número total de observaciones en población. Al elevar al cuadrado cada distancia, logramos que todos los números sean positivos y, al mismo tiempo, asignamos más peso a las desviaciones más grandes (desviación es la distancia entre la media y un valor).

Fórmula para la varianza de población

La fórmula para calcular la varianza es:

Varianza de población

$$\sigma^2 = \frac{\Sigma(x - \mu)^2}{N} = \frac{\Sigma x^2}{N} - \mu^2 \qquad [3\text{-}12]$$

donde:

- σ^2 = varianza de la población
- x = elemento u observación
- μ = media de la población
- N = número total de elementos de la población
- Σ = suma de todos los valores $(x - \mu)^2$, o todos los valores x^2

En la ecuación 3-12, la expresión $\frac{\Sigma(x - \mu)^2}{N}$ es la definición de σ^2. La última expresión, $\frac{\Sigma x^2}{N} - \mu^2$, es *matemáticamente* equivalente a la definición, pero a menudo resulta mucho más conveniente utilizarla si de hecho debemos calcular el valor de σ^2, ya que nos permite no calcular las desviaciones respecto a la media. Sin embargo, cuando los valores de x son grandes y los valores $x - \mu$ pequeños, puede ser más conveniente utilizar la expresión $\frac{\Sigma(x - \mu)^2}{N}$ para calcular σ^2. Antes de poder utilizar esta fórmula en un ejemplo, necesitamos analizar un problema importante referente a la varianza. Al resolver ese problema, aprenderemos qué es la desviación estándar y cómo calcularla. Después, podremos regresar a la varianza en sí.

Las unidades en las que se expresa la varianza ocasionan problemas

Al principio, cuando calculamos el rango, las respuestas se expresaron en las mismas unidades que los datos. (En nuestros ejemplos, las unidades son "pagos de miles de dólares".) Para la varianza, sin embargo, las unidades son el *cuadrado de las unidades* de los datos; por ejemplo, "dólares al cuadrado". Estas unidades no son intuitivamente claras o fáciles de interpretar. Por esto debemos hacer un cambio significativo en la varianza para calcular una medida útil de la desviación que no nos dé problemas con las unidades de medida y, en consecuencia, sea menos confusa. **Esta medida se conoce como la desviación estándar y es la raíz cuadrada de la varianza.** La raíz cuadrada de 100 dólares cuadrados es 10 dólares, puesto que tomamos la raíz cuadrada tanto del valor como de las unidades en que se miden. La desviación estándar, entonces, queda en las mismas unidades que los datos originales.

Desviación estándar de la población

Relación de la desviación estándar y la varianza

La desviación estándar de la población, σ, es simplemente la raíz cuadrada de la varianza de la población. Como la varianza es el promedio de los cuadrados de las distancias de las observaciones a la media, **la desviación estándar es la raíz cuadrada del promedio de los cuadrados de las distancias entre las observaciones y la media**. Mientras que la varianza se expresa con el cuadrado de las unidades utilizadas para medir los datos, la desviación estándar está en las mismas unidades que las que se usaron para medir los datos. La fórmula para la desviación estándar es:

Desviación estándar de la población

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\Sigma(x-\mu)^2}{N}} = \sqrt{\frac{\Sigma x^2}{N} - \mu^2} \qquad [3\text{-}13]$$

donde,

- x = observación
- μ = media de la población
- N = número total de elementos de la población
- Σ = suma de todos los valores $(x - \mu)^2$, o todos los valores x^2
- σ = desviación estándar de la población
- σ^2 = varianza de la población

Utilice la raíz cuadrada positiva

La raíz cuadrada de un número positivo puede ser positiva o negativa, ya que $a^2 = (-a)^2$. Sin embargo, cuando obtenemos la raíz cuadrada de la varianza para calcular la desviación estándar los especialistas en estadística sólo consideran la raíz cuadrada positiva.

Cálculo de la desviación estándar

Para calcular la varianza o la desviación estándar, elaboramos una tabla utilizando todos los elementos de la población. Si tenemos una población de 15 frascos de compuesto producido en un día y probamos cada frasco para determinar la pureza del compuesto, los datos obtenidos podrían ser los de la tabla 3-18. La tabla 3-19 muestra la forma en que se utilizan estos datos para calcular la media ($0.166 = 2.49/15$, suma de los valores de la columna 1 dividida entre N), la desviación de cada valor respecto a la media (columna 3), el cuadrado de la desviación de cada valor respecto a la media (columna 4), y la suma de los cuadrados de las desviaciones. A partir de esto, podemos calcular la varianza, que es del 0.0034% al cuadrado. (En la tabla 3-19 también calculamos σ^2 utilizando la segunda mitad de la ecuación 3-12, $\frac{\Sigma x^2}{N} - \mu^2$. Observe que obtenemos el mismo resultado, pero con menos trabajo, ya que no tenemos que calcular las desviaciones respecto a la media.) Tomando la raíz cuadrada de σ^2, podemos calcular la desviación estándar, 0.058%.

Usos de la desviación estándar

Teorema de Chebyshev

La desviación estándar nos permite determinar, con un buen grado de precisión, dónde están localizados los valores de una distribución de frecuencias con relación a la media. Podemos hacer esto de acuerdo con un teorema establecido por el matemático ruso P. L. Chebyshev (1821-1894). El teorema de Chebyshev establece que independientemente de la forma de la distribución, al menos 75%

Tabla 3-18 Resultados de la prueba de pureza de los compuestos

Porcentaje de impureza observado				
0.04	0.14	0.17	0.19	0.22
0.06	0.14	0.17	0.21	0.24
0.12	0.15	0.18	0.21	0.25

Tabla 3-19

Determinación de la varianza y la desviación estándar del porcentaje de impureza de los compuestos

Observación (x) (1)		Media $\mu = 2.49/15$ (2)		Desviación ($x - \mu$) (3) = (1) − (2)	Desviación al cuadrado $(x - \mu)^2$ (4) = [(1) − (2)]^2	Observación al cuadrado (x^2) (5) = (1)^2
0.04	−	0.166	=	−0.126	0.016	0.0016
0.06	−	0.166	=	−0.106	0.011	0.0036
0.12	−	0.166	=	−0.046	0.002	0.0144
0.14	−	0.166	=	−0.026	0.001	0.0196
0.14	−	0.166	=	−0.026	0.001	0.0196
0.15	−	0.166	=	−0.016	0.000	0.0225
0.17	−	0.166	=	0.004	0.000	0.0289
0.17	−	0.166	=	0.004	0.000	0.0289
0.18	−	0.166	=	0.014	0.000	0.0324
0.19	−	0.166	=	0.024	0.001	0.0361
0.21	−	0.166	=	0.044	0.002	0.0441
0.21	−	0.166	=	0.044	0.002	0.0441
0.22	−	0.166	=	0.054	0.003	0.0484
0.24	−	0.166	=	0.074	0.005	0.0576
0.25	−	0.166	=	0.084	0.007	0.0625
2.49 ← Σx					**0.051** ← $\Sigma(x - \mu)^2$	**0.4643** ← Σx^2

$$\sigma^2 = \frac{\Sigma(x-\mu)^2}{N} \quad [3\text{-}12] \qquad \leftarrow 0 \rightarrow \qquad \sigma^2 = \frac{\Sigma x^2}{N} - \mu^2 \quad [3\text{-}12]$$

$$= \frac{0.051}{15} \qquad\qquad = \frac{0.4643}{15} - (0.166)^2$$

$$= 0.0034 \text{ al cuadrado} \qquad\qquad = 0.0034 \text{ al cuadrado}$$

$$\sigma = \sqrt{\sigma^2} \quad [3\text{-}13]$$

$$= \sqrt{0.0034}$$

$$= 0.058\%$$

de los valores caen dentro de ± 2 desviaciones estándar a partir de la media de la distribución, y al menos 89% de los valores caen dentro de ± 3 desviaciones estándar a partir de la media.

Podemos medir aún con más precisión el porcentaje de observaciones que caen dentro de un rango específico de una curva simétrica con forma de campana, como la mostrada en la figura 3-12. En estos casos, podemos decir que:

1. Aproximadamente 68% de los valores de la población cae dentro de ±1 desviación estándar a partir de la media.
2. Aproximadamente 95% de los valores estará dentro de ±2 desviaciones estándar a partir de la media.

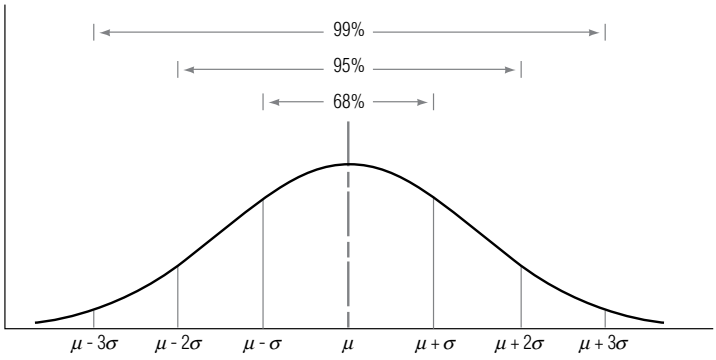


FIGURA 3-12

Localización de las observaciones alrededor de la media para una distribución de frecuencias con forma de campana

3. Aproximadamente 99% de los valores estará en el intervalo que va desde 3 desviaciones estándar a la izquierda de la media hasta 3 desviaciones estándar a la derecha de la media.

Uso del teorema de Chebyshev

A la luz del teorema de Chebyshev, analicemos los datos de la tabla 3-19. En ellos, la impureza media de los 15 frascos de compuesto es 0.166% y la desviación estándar es 0.058%. El teorema de Chebyshev nos dice que al menos el 75% de los valores (11 de nuestros 15 frascos) están entre $0.166 - 2(0.058) = 0.050$ y $0.166 + 2(0.058) = 0.282$. De hecho, 93% de las observaciones (14 de los 15 valores) están realmente en el intervalo. Note que la distribución es razonablemente simétrica y que 93% es muy cercano al 95% teórico para un intervalo de ±2 desviaciones estándar a partir de la media de una curva con forma de campana.

Concepto de resultado estándar

La desviación estándar es útil también para describir cuánto se apartan las observaciones individuales de una distribución de la media de la misma. Una medida que se conoce como *resultado estándar* nos da el número de desviaciones estándar que una observación en particular ocupa por debajo o por encima de la media. Si x simboliza la observación, entonces el resultado estándar calculado a partir de los datos de la población es:

Resultado estándar

$$\text{Resultado estándar de la población} = \frac{x - \mu}{\sigma} \qquad [3\text{-}14]$$

donde,

- x = observación tomada de la población
- μ = media de la población
- σ = desviación estándar de la población

Suponga que observamos un frasco de compuesto que tiene 0.108% de impureza. Como nuestra población tiene una media de 0.166 y una desviación estándar de 0.058, una observación de 0.108 tendría un resultado estándar de -1:

Cálculo del resultado estándar

$$\begin{aligned}\text{Resultado estándar} &= \frac{x - \mu}{\sigma} \qquad [3\text{-}14]\\ &= \frac{0.108 - 0.166}{0.058}\\ &= -\frac{0.058}{0.058}\\ &= -1\end{aligned}$$

Una impureza observada del 0.282% tendría un resultado estándar de $+2$:

$$\begin{aligned}\text{Resultado estándar} &= \frac{x - \mu}{\sigma} \qquad [3\text{-}14]\\ &= \frac{0.282 - 0.166}{0.058}\\ &= \frac{0.116}{0.058}\\ &= 2\end{aligned}$$

Interpretación del resultado estándar

El resultado estándar indica que una impureza del 0.282% se desvía de la media en $2(0.058) = 0.116$ unidades, que es igual a $+2$, en términos de del número de desviaciones estándar alejado de la media.

Cálculo de la varianza y la desviación estándar utilizando datos agrupados

Cálculo de la varianza y de la desviación estándar de datos agrupados

En el ejemplo al principio del capítulo, los datos respecto a las ventas en 100 restaurantes de comida rápida se encuentran agrupados en una distribución de frecuencias. Con esos datos, podemos utilizar las siguientes fórmulas para calcular la varianza y la desviación estándar:

Varianza de datos agrupados

$$\sigma^2 = \frac{\Sigma f(x-\mu)^2}{N} = \frac{\Sigma f x^2}{N} - \mu^2 \qquad [3\text{-}15]$$

Desviación estándar de datos agrupados

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\Sigma f(x-\mu)^2}{N}} = \sqrt{\frac{\Sigma f x^2}{N} - \mu^2} \qquad [3\text{-}16]$$

donde,

- σ^2 = varianza de la población
- σ = desviación estándar de la población
- f = frecuencia de cada una de las clases
- x = punto medio de cada clase
- μ = media de la población
- N = tamaño de la población

La tabla 3-20 muestra cómo aplicar estas ecuaciones para encontrar la varianza y la desviación estándar de las ventas en 100 restaurantes de comida rápida.

Dejamos como ejercicio para el lector curioso verificar que la segunda mitad de la ecuación 3-15, $\frac{\Sigma f x^2}{N} - \mu^2$ da como resultado el mismo valor de σ^2.

Cambio a la varianza y la desviación estándar de una muestra

Ahora estamos listos para calcular las estadísticas de muestra análogas a la varianza de población σ^2 y la desviación estándar de la población, σ. Se trata de la varianza de la muestra s^2 y la desviación estándar de la muestra, s. En la sección siguiente, observará que cambiamos la notación con letras griegas (que denotan parámetros de población) a las latinas correspondientes a las estadísticas de muestras.

Desviación estándar de una muestra

Cálculo de la desviación estándar de una muestra

Para calcular la varianza y la desviación estándar de una muestra, utilizamos las mismas fórmulas de las ecuaciones 3-12 y 3-13, sustituyendo μ con $\bar{x}$ y N con $n - 1$. Las fórmulas tienen el siguiente aspecto:

Varianza de una muestra

$$s^2 = \frac{\Sigma(x-\bar{x})^2}{n-1} = \frac{\Sigma x^2}{n-1} - \frac{n\bar{x}^2}{n-1} \qquad [3\text{-}17]$$

Desviación estándar de una muestra

$$s = \sqrt{s^2} = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}} = \sqrt{\frac{\Sigma x^2}{n-1} - \frac{n\bar{x}^2}{n-1}} \qquad [3\text{-}18]$$

Tabla 3-20

Determinación de la varianza y la desviación estándar de las ventas en 100 restaurantes de comida rápida situados en el distrito del este (miles)

Clase	Punto medio x (1)	Frecuencia f (2)	$f \times x$ (3) = (2) × (1)	Media μ (4)	$x - \mu$ (1) − (4)	$(x - \mu)^2$ [(1) − (4)]^2	$f(x - \mu)^2$ (2) × [(1) − (4)]^2
700- 799	750	4	3,000	1,250	−500	250,000	1,000,000
800- 899	850	7	5,950	1,250	−400	160,000	1,120,000
900- 999	950	8	7,600	1,250	−300	90,000	720,000
1,000-1,099	1,050	10	10,500	1,250	−200	40,000	400,000
1,100-1,199	1,150	12	13,800	1,250	−100	10,000	120,000
1,200-1,299	1,250	17	21,250	1,250	0	0	0
1,300-1,399	1,350	13	17,550	1,250	100	10,000	130,000
1,400-1,499	1,450	10	14,500	1,250	200	40,000	400,000
1,500-1,599	1,550	9	13,950	1,250	300	90,000	810,000
1,600-1,699	1,650	7	11,550	1,250	400	160,000	1,120,000
1,700-1,799	1,750	2	3,500	1,250	500	250,000	500,000
1,800-1,899	1,850	1	1,850	1,250	600	360,000	360,000
		100	**125,000**				**6,680,000**

$$\bar{x} = \frac{\Sigma(f \times x)}{n} \qquad [3\text{-}3]$$

$$= \frac{125{,}000}{100}$$

$$= 1{,}250 \text{ (miles de dólares)} \leftarrow \text{Media}$$

$$\sigma^2 = \frac{\Sigma f(x - \mu)^2}{N} \qquad [3\text{-}15]$$

$$= \frac{6{,}680{,}000}{100}$$

$$= 66{,}800 \text{ (o } 66{,}800 \text{ [miles de dólares]}^2) \leftarrow \text{Varianza}$$

$$\sigma = \sqrt{\sigma^2} \qquad [3\text{-}16]$$

$$= \sqrt{66{,}800}$$

$$= 258.5 \leftarrow \text{Desviación estándar} = \$258{,}500$$

Table 3-21

Determinación de la varianza y la desviación estándar de la muestra de los donativos anuales de Blue Cross-Blue Shield al Hospital de Cumberland (miles)

Observación (x) (1)	Media ($\bar{x}$) (2)	$x - \bar{x}$ (1) − (2)	$(x - \bar{x})^2$ $[(1) - (2)]^2$	x^2 $(1)^2$
863	1,351	−488	238,144	744,769
903	1,351	−448	200,704	815,409
957	1,351	−394	155,236	915,849
1,041	1,351	−310	96,100	1,083,681
1,138	1,351	−213	45,369	1,295,044
1,204	1,351	−147	21,609	1,449,616
1,354	1,351	3	9	1,833,316
1,624	1,351	273	74,529	2,637,376
1,698	1,351	347	120,409	2,883,204
1,745	1,351	394	155,236	3,045,025
1,802	1,351	451	203,401	3,247,204
1,883	1,351	532	283,024	3,545,689
			$\Sigma(x - \bar{x})^2 \rightarrow$ **1,593,770**	**23,496,182** $\leftarrow \Sigma x^2$

$$s^2 = \frac{\Sigma(x - \bar{x})^2}{n - 1} \qquad [3\text{-}17]$$

$$= \frac{1{,}593{,}770}{11}$$

$$= 144{,}888 \text{ (o \$144,888 [miles de dólares]}^2\text{)} \leftarrow \text{Varianza de la muestra}$$

$$s = \sqrt{s^2} \qquad [3\text{-}18]$$

$$= \sqrt{144{,}888}$$

$$= 380.64 \text{ (es decir, \$380,640)} \leftarrow \text{Desviación estándar de la muestra}$$

O

$$s^2 = \frac{\Sigma x^2}{n - 1} - \frac{n\bar{x}^2}{n - 1} \qquad [3\text{-}17]$$

$$= \frac{23{,}496{,}182}{11} - \frac{12(1{,}351)^2}{11}$$

$$= \frac{1{,}593{,}770}{11}$$

$$= 144{,}888$$

donde,

- s^2 = Varianza de la muestra
- s = Desviación estándar de la muestra
- x = Valor de cada una de las n observaciones
- $\bar{x}$ = Media de la muestra
- $n - 1$ = Número de observaciones de la muestra menos 1

Uso de $n - 1$ como denominador

¿Por qué utilizamos $n - 1$ como denominador en lugar de n? Los especialistas en estadística pueden demostrar que si tomamos muchas muestras de una población dada, encontramos la varianza de la muestra (s^2) para cada muestra y promediamos los resultados, este promedio no tiende a igualar el valor de la varianza de la población, σ^2, a menos que usemos $n - 1$ como denominador en nuestros cálculos. En el capítulo 7, se dará la explicación estadística de por qué esto es cierto.

Cálculo de la varianza y la desviación estándar de la muestra para los datos del hospital

Las ecuaciones 3-17 y 3-18 nos permiten encontrar la varianza y la desviación estándar de la muestra de los donativos anuales de Blue Cross-Blue Shield al Hospital de Cumberland que presentamos en la tabla 3-21; observe que ambas mitades de la ecuación 3-17 producen el mismo resultado.

Cálculo de los resultados estándar de la muestra

Igual que utilizamos la desviación estándar de la población para derivar los resultados estándar de la misma, podemos usar la desviación estándar de la muestra para calcular los resultados estándar de la muestra. Estos resultados indican a cuántas desviaciones estándar arriba o abajo de la media de la muestra se encuentra una observación dada. La fórmula adecuada es:

Resultado estándar de una observación de una muestra

$$\text{Resultado estándar de la muestra} = \frac{x - \bar{x}}{s} \qquad [3\text{-}19]$$

donde:

- x = observación tomada de la muestra
- $\bar{x}$ = media de la muestra
- s = desviación estándar de la muestra

En el ejemplo anterior, vemos que la observación 863 corresponde a un resultado estándar de -1.28:

$$\begin{aligned} \text{Resultado estándar de la muestra} &= \frac{x - \bar{x}}{s} \qquad [3\text{-}19] \\ &= \frac{863 - 1{,}351}{380.64} \\ &= \frac{-488}{380.64} \\ &= -1.28 \end{aligned}$$

En esta sección hemos demostrado por qué la desviación estándar es la medida de dispersión que más se utiliza. Podemos usarla para comparar distribuciones y para calcular resultados estándar, que son un elemento importante de la inferencia estadística que analizaremos más adelante. Al igual que la varianza, la desviación estándar toma en cuenta cada observación del conjunto de datos. Sin embargo, la desviación estándar tiene también algunas desventajas. No es fácil calcularla como el rango, y no puede calcularse en distribuciones de extremo abierto. Además, los valores extremos que se encuentren en el conjunto de datos distorsionan el valor de la desviación estándar, aunque en menor grado que en el caso del rango.

SUGERENCIAS Y SUPOSICIONES

Al calcular y usar la desviación estándar se supone que no hay muchos valores demasiado grandes o demasiado pequeños en el conjunto de datos porque se sabe que la desviación estándar usa todos los valores; esos valores extremos distorsionarán la respuesta. Sugerencia: puede evitarse la confusión entre usar N o $n - 1$ como denominador para las muestras y poblaciones si se asocia el valor *más pequeño* ($n - 1$) con el conjunto *más pequeño* (la muestra).

Ejercicios 3.9

Ejercicios de autoevaluación

EA 3-13 Talent, Ltd., una compañía en Hollywood de selección de elenco, está en proceso de elegir un grupo de extras para una película. Las edades de los 20 hombres que se entrevistaron primero son:

50	56	55	49	52	57	56	57	56	59
54	55	61	60	51	59	62	52	54	49

El director de la película quiere hombres cuyas edades se agrupen de manera cercana alrededor de los 55 años. Con sus conocimientos de estadística, el director sugiere que sería aceptable una desviación estándar de 3 años. ¿Califica este grupo de extras?

EA 3-14 En un intento de estimar la demanda potencial futura, la National Motor Company realizó un estudio, en 1988, en el que preguntaba a parejas casadas cuántos automóviles debe tener la familia promedio actual. Para cada pareja, promediaron las repuestas del hombre y la mujer, a fin de obtener la respuesta global de la pareja. Las respuestas se colocaron en una tabla:

Número de autos	0	0.5	1.0	1.5	2.0	2.5
Frecuencia	2	14	23	7	4	2

a) Calcule la varianza y la desviación estándar.
b) Dado que la distribución tiene, casi, forma de campana, en teoría, ¿cuántas observaciones deben caer entre 0.5 y 1.5? ¿Entre 0 y 2? ¿Cuántas caen de hecho en esos intervalos?

Aplicaciones

■ **3-61** La chef en jefe de The Flying Taco acaba de recibir dos docenas de jitomates de su proveedora, pero todavía no los acepta. Sabe por la factura que el peso promedio de un jitomate es 7.5 onzas, pero insiste en que todos tengan un peso uniforme. Aceptará los jitomates sólo si el peso promedio es 7.5 onzas y la desviación estándar es menor que 0.5 onzas. Los pesos de los jitomates son los siguientes:

6.3	7.2	7.3	8.1	7.8	6.8	7.5	7.8	7.2	7.5	8.1	8.2
8.0	7.4	7.6	7.7	7.6	7.4	7.5	8.4	7.4	7.6	6.2	7.4

¿Cuál es la decisión de la chef y por qué?

■ **3-62** Los siguientes datos son una muestra de la tasa de producción diaria de botes de fibra de vidrio de la Hydrosport, Ltd., un fabricante de Miami:

17	21	18	27	17	21	20	22	18	23

El gerente de producción de la compañía siente que una desviación estándar de más de tres botes por día indica variaciones de tasas de producción inaceptables. ¿Deberá preocuparse por las tasas de producción de la planta?

■ **3-63** Un conjunto de 60 observaciones tiene una media de 66.8, una varianza de 12.60 y una forma de distribución desconocida.
a) ¿Entre qué valores deberán caer al menos 75% de las observaciones, de acuerdo con el teorema de Chebyshev?
b) Si la distribución es simétrica y con forma de campana, aproximadamente cuántas observaciones deberán encontrarse en el intervalo 59.7-73.9?
c) Encuentre los resultados estándar para las siguientes observaciones tomadas de la distribución: 61.45, 75.37, 84.65 y 51.50.

■ **3-64** El número de cheques cobrados diariamente en las cinco sucursales del Bank of Orange County durante el mes anterior tuvo la siguiente distribución de frecuencias:

Clase	Frecuencia
0-199	10
200-399	13
400-599	17
600-799	42
800-999	18

Hank Spivey, director de operaciones del banco, sabe que una desviación estándar en el cobro de cheques mayor que 200 cheques diarios ocasiona problemas de personal y de organización en las sucursales, debido a la carga de trabajo dispareja. ¿Deberá preocuparse por la cantidad de empleados que van a utilizar el mes siguiente?

■ **3-65** El consejo directivo del Banco de la Reserva Federal de Estados Unidos ha otorgado permisos a todos los bancos miembros para elevar las tasas de interés 0.5% para todos los depositantes. Las tasas de interés anteriores para cuentas de ahorro eran $5^1/_4$; para certificados de depósito (CD) a un año, $7^1/_2$%; para CD a

18 meses, $8^3/_4$; a dos años, $9^1/_2$; a tres años, $10^1/_2$, y para CD a cinco años, 11%. El presidente del First State Bank desea saber qué características tendrá la nueva distribución de tasas de interés si se le agrega $^1/_2$% a todas las tasas. ¿Cómo se relacionan las nuevas características con las anteriores?

■ **3-66** El administrador de un hospital de Georgia investigó el número de días que 200 pacientes, elegidos al azar, se quedan en el hospital después de una operación. Los datos son:

Frecuencia en el hospital en días	1-3	4-6	7-9	10-12	13-15	16-18	19-21	22-24
Frecuencia	18	90	44	21	9	9	4	5

a) Calcule la desviación estándar y la media.
b) De acuerdo con el teorema de Chebyshev, ¿cuántas estancias habrá entre 0 y 17 días? ¿Cuántas hay realmente en ese intervalo?
c) Debido a que la distribución tiene aproximadamente forma de campana, ¿cuántas estancias entre 0 y 17 días pueden esperarse?

■ **3-67** FundInfo proporciona información a sus suscriptores para permitirles evaluar el desempeño de los fondos de inversión que consideran vehículos de inversión potencial. Un estudio reciente de los fondos cuya meta de inversión establecida era crecimiento e ingreso produjo los siguientes datos de la tasa de retorno anual sobre la inversión total durante los últimos cinco años:

Rendimiento anual (%)	11.0-11.9	12.0-12.9	13.0-13.9	14.0-14.9	15.0-15.9	16.0-16.9	17.0-17.9	18.0-18.9
Frecuencia	2	2	8	10	11	8	3	1

a) Calcule la media, la varianza y la desviación estándar de la tasa de rendimiento anual para esta muestra de 45 fondos de inversión.
b) Según el teorema de Chebyshev, ¿entre qué valores debe caer al menos 75% de las observaciones de la muestra? ¿Qué porcentaje de observaciones caen de hecho en ese intervalo?
c) Dado que la distribución es casi una campana, ¿entre qué valores se esperaría encontrar 68% de las observaciones? ¿Qué porcentaje de las observaciones de hecho caen en ese intervalo?

■ **3-68** Nell Berman, propietario de la Earthbread Bakery, afirmó que el nivel de producción promedio por semana de su empresa fue 11,398 barras de pan, con una varianza de 49,729. Si los datos utilizados para calcular los resultados se recolectaron en el periodo de 32 semanas, ¿durante cuántas semanas estuvo el nivel de producción abajo de 11,175? ¿Y cuántas arriba de 11,844?

■ **3-69** La compañía Creative Illusion Advertising tiene tres oficinas en tres ciudades distintas. Los niveles de salario difieren de un estado a otro. En la oficina de Washington, D.C., el aumento promedio a los salarios durante el año anterior fue $1,500, con una desviación estándar de $400. En la sucursal de Nueva York, el aumento promedio fue $3,760, con una desviación estándar de $622. En Durham N.C., el aumento promedio fue $850, con una desviación estándar de $95. Se entrevistó a tres empleados. El empleado de Washington recibió un aumento de $1,100; el de Nueva York, obtuvo un aumento de $3,200; y el de Durham uno de $500. ¿Cuál de los tres tuvo el menor aumento en relación con la media y la desviación estándar de los aumentos correspondientes a su oficina?

■ **3-70** La American Foods comercializa con fuerza tres de sus productos a nivel nacional. Uno de los objetivos fundamentales de la publicidad de cada producto consiste en lograr que los consumidores reconozcan que American Foods elabora el producto. Para medir qué tan bien cada anuncio logra ese reconocimiento, se le pidió a un grupo de consumidores que identificara lo más rápido posible a la compañía responsable de una larga lista de productos. El primer producto de la American Foods obtuvo un tiempo promedio, antes de ser reconocido, de 2.5 segundos, con una desviación estándar de 0.004 segundos. El segundo producto tuvo un tiempo promedio de 2.8 segundos, con una desviación estándar de 0.006 segundos. E1 tercero, un tiempo promedio de 3.7 segundos, con una desviación estándar de 0.09 segundos. Uno de los encuestados en particular tuvo los siguientes tiempos antes de reconocer la procedencia del producto: 2.495 para el primero, 2.79 para el segundo y 3.90 para el tercero. ¿Para cuál de los productos estuvo el consumidor en cuestión más alejado del desempeño promedio, en unidades de desviación estándar?

■ **3-71** Sid Levinson es un médico especializado en el conocimiento y uso efectivo de medicinas que eliminan el dolor en pacientes gravemente enfermos. Con el fin de saber aproximadamente cuántas enfermeras y personal administrativo debe emplear, ha empezado a registrar el número de pacientes que atiende cada semana. En ese lapso, su administrador registra el número de pacientes gravemente enfermos y el número de pacientes sin mayores problemas. Sid tiene razones para creer que el número de pacientes sin mayores problemas por semana tendría una distribución en forma de campana, si tuviera suficientes datos (es-

to no es cierto para los pacientes gravemente enfermos). Sin embargo, ha recolectado datos sólo durante las cinco últimas semanas.

Pacientes gravemente enfermos	33	50	22	27	48
Pacientes sin mayores problemas	34	31	37	36	27

a) Calcule la media y la varianza para el número de pacientes seriamente enfermos por semana. Utilice el teorema de Chebyshev para encontrar los límites dentro de los cuales deberá caer el "75% central" del número de pacientes gravemente enfermos por semana.
b) Calcule la media, la varianza y la desviación estándar para el número de pacientes sin mayores problemas por semana. ¿Dentro de qué límites deberá caer el "68% central" de estas cifras semanales?

■ **3-72** El inspector de cualquier distrito escolar tiene dos problemas principales: primero, la dificultad de tratar con la directiva escolar elegida y, segundo, la necesidad de estar siempre preparado para buscar un nuevo empleo debido al primer problema. Tom Langley, inspector del distrito escolar 18 no es la excepción. Ha comprendido el valor de entender todas las cifras que aparecen en un presupuesto y de ser capaz de utilizarlas en su provecho. Este año, la junta directiva sugirió un presupuesto de investigación de medios de $350,000. Por experiencias anteriores, Tom sabe que el gasto real siempre sobrepasa al presupuesto solicitado, y el excedente tiene una media de $40,000 y una varianza de 100,000,000 de dólares cuadrados. Tom aprendió el teorema de Chebyshev cuando estuvo en la universidad, y piensa que podría serle útil para encontrar un intervalo de valores dentro del cual se encuentre el gasto real 75% del tiempo en los años en que la propuesta de presupuesto sea igual a la de este año. Haga un favor a Tom y encuentre ese intervalo.

■ **3-73** Bea Reele, una prestigiada sicóloga clínica, tiene registros muy precisos sobre todos sus pacientes. A partir de los datos, ha creado cuatro categorías dentro de las cuales puede colocar a todos sus pacientes: niños, adultos jóvenes, adultos y ancianos. Para cada categoría, la sicóloga ha calculado el Coeficiente Intelectual (CI) medio y la varianza de los coeficientes intelectuales dentro de la categoría. Las cifras que obtuvo se presentan en la tabla siguiente. Durante cierto día Bea atendió a cuatro pacientes (uno de cada categoría) y sus CI fueron: niño, 90; adulto joven, 92; adulto, 100, y anciano, 98. ¿Cuál de los pacientes tiene el CI más alejado de la media, en unidades de desviación estándar, correspondiente a esa categoría en particular?

Categoría	CI medio	Varianza de CI
Niño	110	81
Adulto joven	90	64
Adulto	95	49
Anciano	90	121

Soluciones a los ejercicios de autoevaluación

EA 3-13

x	$x - \bar{x}$	$(x - \bar{x})^2$	x	$x - \bar{x}$	$(x - \bar{x})^2$
50	−5.2	27.04	54	−1.2	1.44
56	0.8	0.64	55	−0.2	0.04
55	−0.2	0.04	61	5.8	33.64
49	−6.2	38.44	60	4.8	23.04
52	−3.2	10.24	51	−4.2	17.64
57	1.8	3.24	59	3.8	14.44
56	0.8	0.64	62	6.8	46.24
57	1.8	3.24	52	−3.2	10.24
56	0.8	0.64	54	−1.2	1.44
59	3.8	14.44	49	−6.2	38.44
			1,104		**285.20**

$$\bar{x} = \frac{\Sigma x}{n} = \frac{1{,}104}{20} = 55.2 \text{ años, que es cercano a los 55 años deseados}$$

$$s = \sqrt{\frac{\Sigma (x - \bar{x})^2}{n - 1}} = \sqrt{\frac{285.20}{19}} = 3.874 \text{ años, que muestra más variabilidad que la deseada}$$

EA 3-14 a)

# de autos x	Frecuencia f	$f \times x$	$x - \bar{x}$	$(x - \bar{x})^2$	$f(x - \bar{x})^2$
0	2	0	−1.0288	1.0585	2.1170
0.5	14	7	−0.5288	0.2797	3.9155
1	23	23	−0.0288	0.0008	0.0191
1.5	7	10.5	0.4712	0.2220	1.5539
2	4	8	0.9712	0.9431	3.7726
2.5	2	5	1.4712	2.1643	4.3286
	52	**53.5**			**15.7067**

$$\bar{x} = \frac{\Sigma x}{n} = \frac{53.5}{52} = 1.0288 \text{ autos}$$

$$s^2 = \frac{\Sigma f(x - \bar{x})^2}{n - 1} = \frac{15.707}{51} = 0.3080 \quad \text{así} \quad s = \sqrt{0.3080} = 0.55 \text{ autos}$$

b) (0.5, 1.5) es aproximadamente $\bar{x} \pm s$ entonces, cerca del 68% de los datos, o 0.68(52) = 35.36 observaciones deben estar en este intervalo. De hecho, 44 observaciones están ahí.
(0, 2) es aproximadamente $\bar{x} \pm 2s$, entonces alrededor del 95% de los datos, o 0.95(52) = 49.4 observaciones deben estar en este intervalo. De hecho, 50 observaciones caen en él.

3.10 Dispersión relativa: el coeficiente de variación

La desviación estándar es una medida *absoluta* de la dispersión que expresa la variación en las mismas unidades que los datos originales. Los donativos anuales de Blue Cross-Blue Shield al Hospital de Cumberland (tabla 3-21) tienen una desviación estándar de $380,640, y los que hacen al Hospital de Valley Falls (tabla 3-16), tienen una desviación estándar de $57,390 (que puede usted calcular). ¿Podemos comparar los valores de estas dos desviaciones estándar? Desafortunadamente, la respuesta es no.

Defectos de la desviación estándar

La desviación estándar no puede ser la única base para la comparación de dos distribuciones. Si tenemos una desviación estándar de 10 y una media de 5, los valores varían en una cantidad que es el doble de la media. Si, por otro lado, tenemos una desviación estándar de 10 y una media de 5,000, la variación relativa a la media es insignificante. En consecuencia, no podemos conocer la dispersión de un conjunto de datos hasta que conocemos su desviación estándar, su media y cómo se compara la desviación estándar con la media.

El coeficiente de variación, una medida relativa

Lo que necesitamos es una medida *relativa* que nos proporcione una estimación de la magnitud de la desviación respecto a la magnitud de la media. El *coeficiente de variación* es una de estas medidas relativas de dispersión. Relaciona la desviación estándar y la media, expresando la desviación estándar como porcentaje de la media. La unidad de medida, entonces, es "porcentaje", en lugar de las unidades de los datos originales. Para una población, la fórmula para el coeficiente de variación es:

Coeficiente de variación

$$\text{Coeficiente de variación de la población} = \frac{\sigma}{\mu}(100) \qquad [3\text{-}20]$$

(σ: Desviación estándar de la población; μ: Media de la población)

Para utilizar esta fórmula en un ejemplo, podemos suponer que cada día el técnico A del laboratorio realiza un promedio de 40 análisis con una desviación estándar de 5. El técnico B efectúa un promedio de 160 análisis diarios con una desviación estándar de 15. ¿Cuál de los dos técnicos muestra menos variabilidad?

A primera vista, parece que el técnico B tiene una variación en su producción tres veces mayor que el técnico A. Pero B realiza sus análisis con una rapidez cuatro veces mayor que A. Tomando en cuenta toda esta información, podemos calcular el coeficiente de variación para ambos técnicos:

$$\text{Coeficiente de variación} = \frac{\sigma}{\mu}(100) \qquad [3\text{-}20]$$

Cálculo del coeficiente de variación

$$= \frac{5}{40}(100)$$

$$= 12.5\% \leftarrow \text{Para el técnico A}$$

y

$$\text{Coeficiente de variación} = \frac{15}{160}(100)$$

$$= 94\% \leftarrow \text{Para el técnico B}$$

Así, tenemos que el técnico B, quien tiene una variación *absoluta* mayor que la del técnico A, tiene una variación *relativa* menor que la de A, debido a que la media de producción de B es mucho mayor que la de A.

Uso de la computadora para calcular medidas de tendencia central y de variabilidad

Para conjuntos grandes de datos, utilizamos la computadora para calcular nuestras medidas de tendencia central y de variabilidad. En la figura 3-13, utilizamos el sistema Minitab para calcular algunas de las estadísticas sumarias para los datos de calificaciones dados en el apéndice 10. Las estadísticas se muestran para cada sección, así como para el curso completo. En la figura 3-14 utilizamos Minitab para calcular varias medidas de tendencia central y de variabilidad para los datos sobre ganancias del apéndice 11. Las estadísticas se dan para las 224 compañías juntas, y también se desglosan por bolsa de valores (1 = OTC, 2 = ASE, 3 = NYSE). La estadística MEDREC (TRMEAN, *trimed mean*) es una "media recortada", es decir, una media calculada sin tomar en cuenta el 5% de los datos más altos ni el 5% de los datos más bajos. Esto ayuda a disminuir la distorsión ocasionada por los valores extremos que tanto afectan a la media aritmética.

SUGERENCIAS Y SUPOSICIONES

El concepto y la utilidad del coeficiente de variación son evidentes si se intenta comparar a hombres con sobrepeso y mujeres con sobrepeso. Suponga que un grupo de hombres y mujeres tiene un sobrepeso de 20 libras. Esas 20 libras no son una buena medida del peso excesivo. El peso promedio para los hombres es cerca de 160 libras, mientras que el peso promedio para las mujeres es alrededor de 120 libras. Con un cociente sencillo se puede ver que las mujeres tienen 20/120, es decir 16.7%, de sobrepeso, y el de los hombres corresponde a 20/160, o sea cerca del 12.5%. Aunque el coeficiente de variación es un poco más complejo que el cociente del ejemplo, el concepto es el mismo: se usa para comparar la cantidad de variación en grupos de datos que tienen medias diferentes. Advertencia: no compare la dispersión en los conjuntos de datos usando las desviaciones estándar, a menos que las medias sean parecidas.

Ejercicios 3.10

Ejercicios de autoevaluación

EA 3-15 Basart Electronics piensa emplear uno de dos programas de capacitación. Se capacitó a dos grupos para la misma tarea. El grupo 1 recibió el programa A; el grupo 2, el B. Para el primer grupo, los tiempos requeridos para capacitar a los empleados tuvieron un promedio de 32.11 horas y una varianza de 68.09. En el segundo grupo, el promedio fue 19.75 horas y la varianza fue 71.14. ¿Qué programa de capacitación tiene menos variabilidad relativa en su desempeño?

EA 3-16 Southeastern Stereos, un distribuidor, deseaba convertirse en el proveedor de tres tiendas, pero los faltantes en el inventario lo forzaron a seleccionar sólo uno. El gerente de crédito de Southeastern está evaluando los registros de crédito de estas tres tiendas. En los últimos 5 años, Las cuentas por cobrar de las

Estadística descriptiva

Variable	SECCIÓN	N	Media	Mediana	MedRec	DesvEst	MediaSE	Mín	Máx	Q1	Q3
EXAM1		199	50.22	50.00	50.26	9.49	0.67	21.00	73.00	44.00	57.00
	1	27	47.15	47.00	47.32	10.86	2.09	21.00	69.00	40.00	55.00
	2	46	50.83	50.50	50.83	10.61	1.56	30.00	73.00	43.00	58.25
	3	37	53.19	55.00	53.39	8.98	1.48	35.00	68.00	47.50	60.00
	4	26	50.77	51.50	50.87	8.75	1.72	31.00	68.00	44.75	57.00
	5	36	49.47	48.50	49.16	8.16	1.36	35.00	72.00	44.00	54.75
	6	27	48.67	50.00	48.56	8.44	1.62	34.00	66.00	41.00	54.00
EXAM2		199	56.89	59.00	57.71	10.71	0.76	16.00	73.00	51.00	65.00
	1	27	53.30	56..00	54.20	13.59	2.61	16.00	68.00	49.00	63.00
	2	46	58.26	59.00	59.00	10.84	1.60	24.00	73.00	53.75	67.50
	3	37	60.51	62.00	60.76	7.60	1.25	44.00	72.00	55.00	66.00
	4	26	59.38	59.00	59.46	6.44	1.26	45.00	72.00	55.50	64.25
	5	36	55.94	57.00	56.69	11.44	1.91	25.00	72.00	48.50	65.00
	6	27	52.07	54.00	52.44	11.09	2.13	30.00	65.00	41.00	62.00
TAREA		199	108.60	113.00	110.28	19.01	1.35	13.00	135.00	101.00	121.00
	1	27	109.07	112.00	111.16	20.51	3.95	32.00	134.00	106.00	121.00
	2	46	112.52	116.50	113.90	17.64	2.60	56.00	135.00	107.00	124.00
	3	37	111.78	114.00	113.73	16.80	2.76	35.00	131.00	106.50	122.00
	4	26	104.58	108.00	105.42	15.04	2.95	62.00	127.00	99.00	115.00
	5	36	107.36	114.00	110.19	24.34	4.06	13.00	133.00	98.25	124.00
	6	27	102.59	105.00	102.76	17.03	3.28	74.00	127.00	85.00	120.00
FINAL		199	45.28	45.00	45.53	10.01	0.71	13.00	74.00	39.00	52.00
	1	27	45.74	45.00	46.28	10.68	2.06	14.00	64.00	41.00	53.00
	2	46	44.76	44.00	44.98	11.90	1.75	13.00	74.00	37.75	52.25
	3	37	49.08	49.00	49.27	7.37	1.21	34.00	63.00	43.00	55.00
	4	26	44.92	45.00	44.96	8.06	1.58	29.00	60.00	37.75	51.25
	5	36	44.33	44.00	44.22	10.37	1.73	25.00	65.00	36.00	50.75
	6	27	42.11	44.00	42.60	9.44	1.82	17.00	55.00	38.00	50.00
TOTAL		199	68.57	69.51	68.95	11.24	0.80	22.01	98.11	62.69	75.97
	1	27	67.10	67.00	68.11	13.62	2.62	22.00	87.05	62.87	76.08
	2	46	69.39	71.30	69.62	12.50	1.84	37.79	98.11	63.91	76.41
	3	37	72.82	73.18	73.08	8.86	1.46	53.38	88.21	68.96	80.44
	4	26	68.60	69.38	68.90	8.08	1.59	49.05	81.06	65.50	73.59
	5	36	67.43	65.51	67.67	11.82	1.97	40.91	92.34	60.19	76.23
	6	27	64.30	64.90	64.49	9.85	1.90	43.89	79.85	58.59	75.12

FIGURA 3-13 Salida de Minitab que muestra el resumen estadístico para las calificaciones del curso

Estadística descriptiva

Variable	MERCADO	N	Media	Mediana	MediaTrim	DesvEst	MediaSEM	Mín	Máx	Q1	Q2
LQ89		224	0.2105	0.1300	0.2139	0.8916	0.0556	-5.4500	5.2300	-0.0075	-0.4400
	1	111	0.0766	0.1100	0.1070	0.5110	0.0485	-3.7500	1.2200	-0.0200	0.2600
	2	38	0.199	0.045	0.083	0.837	0.136	-0.560	4.740	-0.085	0.292
	3	75	0.415	0.440	0.459	1.130	0.130	-5.450	5.230	0.070	0.810

FIGURA 3-14 Salida de Minitab que muestra el resumen estadístico para los datos de ingresos

tiendas han sido sobresalientes por los siguientes números de días. El gerente siente que es importante la consistencia, además del promedio menor. Con base en la dispersión relativa, ¿qué tienda sería el mejor cliente?

Lee	62.2	61.8	63.4	63.0	61.7
Forrest	62.5	61.9	62.8	63.0	60.7
Davis	62.0	61.9	63.0	63.9	61.5

Aplicaciones

■ **3-74** El peso de los integrantes del equipo de fútbol americano profesional Baltimore Bullets tiene media de 224 libras con desviación estándar de 18 libras, mientras que los mismos datos correspondientes a su oponente del próximo domingo, los Trailblazers de Chicago, son 195 y 12, respectivamente. ¿Cuál de los dos equipos muestra mayor dispersión relativa respecto al peso de sus integrantes?

■ **3-75** Una universidad ha decidido probar tres nuevos tipos de focos. Tienen tres habitaciones idénticas para realizar el experimento. El foco 1 tiene una vida promedio de 1,470 horas y una varianza de 156. El foco 2 tiene una vida promedio de 1,400 horas y una varianza de 81. La vida promedio del foco 3 es 1,350 horas con una desviación estándar de 6 horas. Clasifique los focos en términos de la variabilidad relativa. ¿Cuál es el mejor?

■ **3-76** La edad de los estudiantes regulares que acuden a un curso en los turnos matutino y vespertino del nivel licenciatura de la Universidad Central se describe en las siguientes dos muestras:

Turno matutino	23	29	27	22	24	21	25	26	27	24
Turno vespertino	27	34	30	29	28	30	34	35	28	29

Si la homogeneidad de la clase es un factor positivo en el aprendizaje, utilice una medida de variabilidad relativa para sugerir en cuál de los dos grupos será más fácil enseñar.

■ **3-77** Existe cierto número de medidas posibles del desempeño de ventas, incluyendo qué tan coherente es un vendedor en el logro de los objetivos de ventas establecidos. Los datos que presentamos a continuación son un registro del porcentaje de los objetivos logrados por tres vendedores durante los 5 años pasados.

Patricia	88	68	89	92	103
John	76	88	90	86	79
Frank	104	88	118	88	123

a) ¿Cuál vendedor es más coherente?
b) Comente sobre lo adecuado de utilizar una medida coherente junto con porcentajes de objetivos de ventas logrados para evaluar el desempeño de ventas.
c) ¿Puede usted sugerir una medida alternativa más apropiada de consistencia?

■ **3-78** La junta directiva de la empresa Gothic Products está considerando adquirir una o dos compañías y examinando minuciosamente la administración de cada compañía, con el fin de hacer una transacción lo menos riesgosa posible. Durante los últimos 5 años, la primera de las compañías tuvo una recuperación promedio de lo invertido del 28.0%, con una desviación estándar del 5.3%. La otra compañía tuvo una recuperación promedio de lo invertido del 37.8%, con una desviación estándar del 4.8%. Si consideramos riesgoso asociarse con una compañía que tenga una alta dispersión relativa en la recuperación, ¿cuál de estas dos compañías ha seguido una estrategia más riesgosa?

■ **3-79** Un laboratorio médico, que provee medicamentos predosificados a un hospital, utiliza diferentes máquinas para los medicamentos que requieren cantidades de dosis diferentes. Una máquina, diseñada para producir dosis de 100 cc, tiene como dosis media 100 cc, con una desviación estándar de 5.2 cc. Otra máquina produce cantidades promediadas de 180 cc de medicamento y tiene una desviación estándar de 8.6 cc. ¿Cuál de las máquinas tiene la menor precisión desde el punto de vista de la dispersión relativa?

■ **3-80** HumanPower, una agencia de empleos temporales, ha probado las habilidades para la captura de datos de muchas personas. Infotech necesita un capturista rápido y consistente. HumanPower revisa los registros de velocidad de 4 empleados con los siguientes datos en términos del número de entradas correctas por minuto. ¿Qué empleado es el mejor para Infotech, según la dispersión relativa?

John	63	66	68	62	69	72
Jeff	68	67	66	67	69	
Mary	62	79	75	59	72	84
Tammy	64	68	58	57	59	

■ **3-81** La compañía Wyatt Seed vende tres categorías de semilla de maíz Early White Sugar, que se diferencian entre sí por el nivel de consistencia de sus germinaciones. El laboratorio de pruebas de semillas del estado tiene una muestra de cada categoría y los resultados de las pruebas acerca del número de semillas que germinan, de un paquete de 100, son los siguientes:

Categoría I (Regular)	88	91	92	89	79
Categoría II (Extra)	87	92	88	90	92
Categoría III (Super)	90	89	79	93	88

¿Tiene sentido la clasificación de semillas qué hace la Wyatt Seed?

■ **3-82** La compañía de electrodomésticos Sunray Appliance acaba de terminar un estudio de la configuración posible de tres líneas de ensamble para producir el tostador doble que más ventas le reporta. La configuración I consume un tiempo medio de 34.8 minutos para construir un aparato, con una desviación estándar de 4.8 minutos. La configuración II produce un tostador en un tiempo medio de 25.5 minutos, con una desviación estándar de 7.5 minutos. La configuración III produce un aparato en un tiempo medio de 37.5 minutos, con una desviación estándar de 3.8 minutos. ¿Qué configuración de línea de ensamble tiene la menor variación relativa en el tiempo que le lleva construir un tostador?

Soluciones a los ejercicios de autoevaluación

EA 3-15 Programa A: $CV = \frac{\sigma}{\mu}(100) = \frac{\sqrt{68.09}(100)}{32.11} = 25.7\%$

Programa B: $CV = \frac{\sigma}{\mu}(100) = \frac{\sqrt{71.14}(100)}{19.75} = 42.7\%$

El programa A tiene menos variabilidad relativa.

EA 3-16 Lee: $\bar{x} = 62.42$ $s = 0{,}7497$ $CV = (s/\bar{x})(100) = \frac{0.7497(100)}{62.42} = 1.20\%$

Forrest: $\bar{x} = 62.18$ $s = 0.9257$ $CV = (s/\bar{x})(100) = \frac{0.9257(100)}{62.18} = 1.49\%$

Davis: $\bar{x} = 62.46$ $s = 0.9762$ $CV = (s/\bar{x})(100) = \frac{0.9762(100)}{62.46} = 1.56\%$

Con base en la dispersión relativa, Lee sería el mejor cliente, pero en realidad no hay mucha diferencia entre los tres.

3.11 Análisis exploratorio de datos (AED)

Las técnicas en esta sección nos permiten revisar muchos datos y resumirlos con rapidez usando algo tan sencillo como aritmética básica y unos cuantos diagramas simples.

En cierto sentido, es justo lo que hemos estado haciendo en los capítulos 2 y 3, pero en cada situación, al construir la distribución de frecuencias y el histograma, se perdió parte de la información. Observe la distribución de frecuencias de la tabla 3-22 de las calificaciones del examen parcial. A partir de esta distribución es imposible saber cómo se distribuyen las calificaciones entre 70-79, a menos que se tenga el conjunto de datos originales.

Una de las técnicas más útiles del análisis exploratorio, la *gráfica de tallo y hoja*, resuelve este problema de manera muy efectiva. Proporciona el *orden de clasificación* de los elementos del conjunto de datos *y* la forma de la distribución.

Tabla 3-22

Calificaciones en el examen parcial con la distribución de frecuencias

79	78	78	67	76	87	85	73	66
99	84	72	66	57	94	84	72	63
51	48	50	61	71	82	93	100	89

	Frecuencia
40-49	1
50-59	3
60-69	5
70-79	8
80-89	6
90-99	3
>99	1

Para producir una gráfica de tallo y hoja para los datos de la tabla 3-22, se hace una lista vertical de los *tallos* (los primeros dígitos de cada elemento de los datos) como sigue:

4
5
6
7
8
9
10

Después se dibuja una línea vertical a la derecha de estos tallos y se listan las *hojas* (el siguiente dígito para cada tallo) a la derecha de la línea en el orden en que aparecen en el conjunto de datos original.

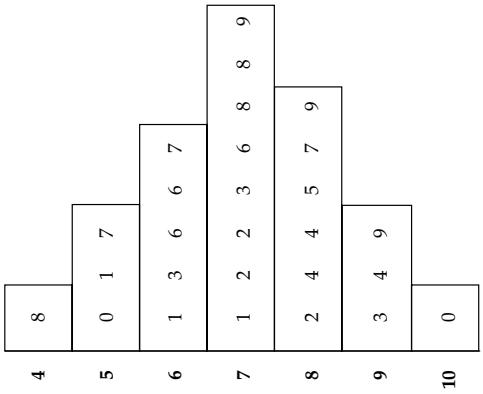
4	8
5	7 1 0
6	7 6 6 3 1
7	9 8 8 6 3 2 2 1
8	7 5 4 4 2 9
9	9 4 3
10	0

Por último se ordenan todas las hojas en cada renglón en orden de clasificación.

4	8
5	0 1 7
6	1 3 6 6 7
7	1 2 2 3 6 8 8 9
8	2 4 4 5 7 9
9	3 4 9
10	0

Cada renglón en la gráfica de tallo y hoja obtenida corresponde a un tallo, y cada valor en ese tallo es una hoja. El renglón **9** | 3 4 9, significa que hay tres elementos en este conjunto de datos que comienzan con 9 (93, 94 y 99).

Si se gira la página 90° en sentido contrario a las manecillas del reloj, se obtiene algo que se parece a los histogramas del capítulo 2.



Alternativas para efectuar análisis exploratorio

Los paquetes de computación que más se utilizan para análisis estadístico tienen la capacidad de procesar un AED. En la figura 3-15 se dan los resultados obtenidos con el paquete SPSS, utilizado para llevar a cabo un análisis exploratorio elemental de datos acerca de los telares para alfombra que vimos en el capitulo 2. Examinaremos brevemente el resultado; si desea saber más acerca del AED, la bibliografía al final del libro proporciona varias referencias.

```
  ILUSTRACIÓN DEL USO DE SPSS PARA HACER ANÁLISIS EXPLORATORIO
                    DE DATOS MONOVARIADA
VARIABLE = YDS      PRODUCCIÓN POR TELAR EN YARDAS
                         MOMENTOS
          N                 30          PESOS    SUMADOS30
          MEDIA          16.0367          SUMA        481.1
          DES EST       0.411459      VARIANZA     0.169299
          SESGO         0.345475      CURTOSIS     -0.10233
          USS            7720.15           CSS      4.90967
          CV             2.56574     MEDIA EST    0.0751219
          MEDIA = 0      213.475      PROB> |T|       0.0001
          CATEG            232.5      PROB> |S|       0.0001
          NUM ~= 0            30
          W:NORMAL      0.969853        PROB<W        0.571
                    CUARTILES (DEF = 4)
           100% MAX        16.9          99%         16.9
            75% Q3         16.3          95%       16.845
            50% MED          16          90%        16.78
            25% Ql       15.775          10%         15.6
             0% MIN        15.2           5%        15.31
                                          1%         15.2
            RANGO           1.7
            Q3-Ql      0.524988
            MODO            15.9
                        EXTREMOS
                     MENOR              MAYOR
                      15.2               16.4
                      15.4               16.6
                      15.6               16.8
                      15.6               16.8
                      15.6               16.9
          TALLO HOJA                #     GRÁFICA DE CAJA [INSERTAR]
           168  000                 3                 |
           166  0                   1                 |
           164  00                  2                 |
           162  00000               5              +-----+
           160  00000               5              *--+--*
           158  0000000             7              |     |
           156  00000               5              +-----+
           154  0                   1                 |
           152  0                   1                 |
           150
              ----+----+----+----+
              MULTIPLIQUE TALLO HOJA POR 10**-01
```

FIGURA 3-15

Análisis exploratorio de datos acerca de los telares para alfombra del capítulo 2, utilizando el paquete de computadora SPSS

En la primera sección de resultados (con encabezado "momentos") se tienen media, desviación estándar y las medidas numéricas del sesgo y la curtosis de los datos; como lo vimos en los capítulos 2 y 3, estas cantidades nos dan información acerca de la *forma* de los datos.

Cuartiles, rangos y percentiles

La siguiente sección de resultados (con encabezado "cuartiles") se refiere a los cuartiles y a diferentes rangos, así como a distintos percentiles que perfilan los extremos superior (99%, 95%, 90%) e inferior (10%, 5%, 1%) de los datos. Así, con un AED no sólo se identifica el centro de los datos, sino que también se atrae la atención sobre los valores no centrales y atípicos que se presentan en el conjunto de datos. A menudo, un examen más detallado de estos puntos "exteriores" mostrará que en realidad no deben estar en el conjunto de datos (quizá fueron errores de registro). Ya hemos visto cómo tales puntos exteriores distorsionan la media de las muestras.

Representación gráfica de los datos

SPSS entonces nos da varias representaciones gráficas de los datos. Las representaciones de "tallo y hoja" son parecidas a los histogramas, pero en éstas se muestran simultáneamente todos los valores de los datos mientras se les agrupa. En consecuencia, poseen la ventaja del histograma, en el sentido de resumir los datos, sin perder detalle. Las "gráficas de caja" nos proporcionan una representación gráfica de la mediana (la recta horizontal en medio de la figura 3-15), los cuartiles (las rectas horizontales superior e inferior de la caja de la figura 3-15) y los extremos (los "bigotes" que salen de la caja). Tal vez sea útil pensar en una gráfica de caja como en el esqueleto de una distribución de frecuencias.

En las figuras 3-16 y 3-17 se muestra algo del AED que se puede hacer con Minitab. En la figura 3-16 utilizamos Minitab para generar una gráfica de tallo y hoja de los datos de ingresos del apéndice 11.

La figura 3-17 muestra gráficas de caja de los datos de ingresos como un todo y divididos por bolsa de valores. De nuevo, el AED llama nuestra atención hacia los puntos exteriores, las observaciones alejadas del centro de la distribución. En las gráficas de caja, estas observaciones exteriores se representan con la letra "O". Un examen más minucioso del conjunto de datos nos muestra que para los dos puntos exteriores más extremos, las compañías descontinuaron algunas de sus operaciones; en un caso (Airgas, Inc.), recibió una suma grande por la venta de sus operaciones descontinuadas, y en el otro (Monarch Capital Corp.), se incurre en un alto costo por el cierre de las operaciones descontinuadas. Debido a estos factores extraordinarios, los dos datos puntuales podrían excluirse del análisis posterior del conjunto de datos.

Presentación de tallo y hojas con caracteres

```
Tallo y hojas de UT89         N = 224
Unidad de hoja = 0.10

     1   -5  4
     1   -4
     3   -3  76
     3   -2
     4   -1  6
    57   -0  876655555433322222222211111111111000000000000000000
  (150)   0  000000000000000000000000000000000000000011111111111111111111111111111+
    17    1  00111111223349
     3    2  3
     2    3
     2    4  7
     1    5  2
```

FIGURA 3-16 Gráfica de tallo y hojas de Minitab para los ingresos del último trimestre de 1989

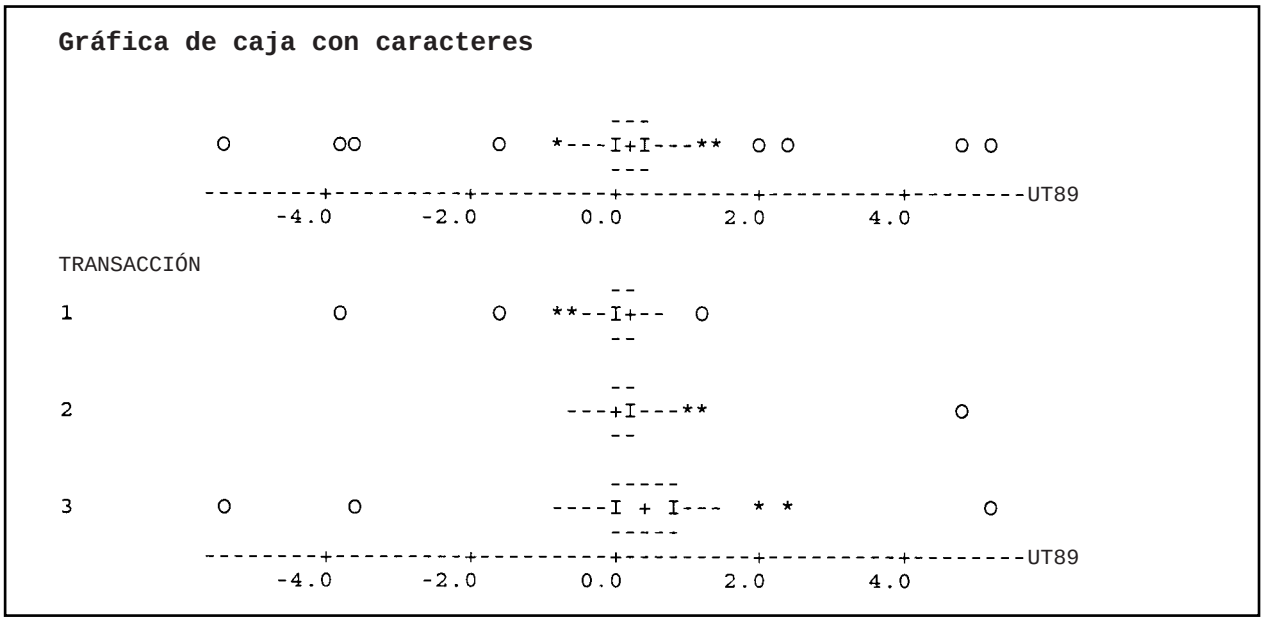


FIGURA 3-17 Gráfica de caja de Minitab para los datos de ingresos

Estadística en el trabajo

Loveland Computers

Caso 3: tendencia central y dispersión "No está mal para unos cuantos días de trabajo, Lee", felicitó el tío Walter a su nuevo asistente mientras echaba un vistazo a las 12 páginas de tablas, diagramas y gráficas. Para Lee, el lunes por la mañana había llegado demasiado aprisa.

"Bueno, Nunc", respondió Lee con esa familiaridad que sólo es posible en una empresa familiar, "me costó algunas desveladas. Pero ordené las cosas de modo que no tengamos que pasar por todo esto en el futuro. Archivé en discos con un mismo formato todos los datos viejos y dejé los de los últimos tres años en el disco duro. Pero lo más importante es que he diseñado algunos formatos comunes para informes de cada una de las líneas de producción, de modo que los datos sean registrados de manera coherente de aquí en adelante. Y con la hoja de cálculo 3D puedo juntar fácilmente toda la información y darte un informe por mes o por trimestre". Alentado por la singular audiencia, Lee pasó a la última página de su informe y mostró a su tío un sencillo diagrama de pastel. "He aquí la belleza del negocio: puedes mostrarles a esos neoyorquinos que tu margen promedio de recuperación (ya sabes, las entradas menos el costo de los bienes vendidos) es del 28%. Eso les va a impresionar."

"Puede que sí, puede que no", comentó Gracia Delaguardia, la socia de Walter Azko, que acababa de entrar a la oficina. Si Walter era reconocido por su encanto y su "astucia callejera", Gracia se había ganado el título de "cerebro" de la compañía. "Tal vez estés mezclando peras con manzanas. Algunas de las PC de baja velocidad ya no tienen ese margen de recuperación. La ganancia es un poco baja, pero al menos se puede predecir. Con las nuevas tecnologías, tenemos un enorme margen en nuestros productos 'novedosos', pero hay otros productos en los que tuvimos que bajar los precios para deshacernos de ellos. Recordarás nuestra primera "portátil" que pesaba más de 22 kilogramos, Walt."

"Estoy intentando olvidarla", respondió el jefe rápidamente. "Pero, Lee, Gracia tiene razón. ¿No crees que tienes que separar los nuevos productos —digamos, productos que tengan menos de seis meses a la venta— de las líneas establecidas? Revisa si el margen de recuperación se ve diferente y si esto se reproduce en todas partes como dice Gracia. Me marcho al aeropuerto a recoger a los inversionistas. A ver qué puedes obtener en lo que regreso."

Preguntas de estudio: El programa de hoja de cálculo que utilizó Lee tiene muchas funciones estadísticas integradas. ¿Cuáles debe usar Lee para responder a las interrogantes sobre los márgenes de recuperación? ¿De qué manera deben presentarse los datos y cómo influirá esto en los nuevos inversionistas para tomar decisiones? ¿Qué limitaciones existen al suponer una distribución con forma de campana para los datos de "porcentaje"?

Ejercicio de base de datos computacional

HH Industries

De lo que Laurel podía ver, las tendencias identificadas en los histogramas que había preparado reflejaban varias posibilidades:

1. El número de clientes que compra a HH Industries está aumentando y sus adquisiciones iniciales son relativamente pequeñas. Con un "cuidado" adecuado, estos clientes deberán comprar cantidades cada vez más grandes, conforme aumente su confianza en la calidad y servicio de la compañía.
2. Los clientes habituales de HH Industries empiezan a reconocer el gasto que implica mantener inventarios grandes. En consecuencia, sus pedidos serán más frecuentes y por cantidades menores que antes.
3. Las empresas constructoras grandes o las que manejan desechos que por tradición han mantenido sus propias flotillas de equipo, pueden tender cada vez más hacia la contratación de pequeños talleres para la reparación de sus averías.

Además, podría estar presente una tendencia estacional, pues las condiciones climáticas adversas del invierno podrían ocasionar retrasos en las construcciones (este concepto será abordado en un capítulo posterior).

Stan Hutchings, vicepresidente de ventas, verificó que todas estas tendencias fueran posibilidades bastante reales.

"De hecho", reconoció, "tal vez podamos lanzar algunas promociones para aprovechar esta información. Pero incluso con la disminución de las cifras de dinero por pedido, estoy realmente convencido de que las ventas totales seguirán aumentando; al menos eso es lo que las 'cifras diarias' parecen indicar. Después de todo, es lo que realmente importa, ¡las ventas totales!"

De regreso en su oficina, Laurel se puso a pensar en la filosofía de Stan sobre las "ventas totales". Era cierto, las ventas totales de la compañía eran importantes, pero ella sabía que cada centro de distribución desempeña un papel clave en la salud global de la compañía. Necesitaba saber si los tres centros estaban haciendo lo suyo. Además, Laurel sentía curiosidad sobre si la tendencia que había identificado se veía reflejada en cada centro de distribución, o si la mayor parte de las ventas hechas en la sucursal de Florida oscurecían cualquier información importante proveniente de los almacenes de Arizona y Pennsylvania.

1. Calcule la media, la mediana y la moda de los datos trimestrales con respecto al número y tamaño promedio de los pedidos, tomando en cuenta los datos del ejercicio correspondiente del capítulo 2. ¿Los resultados apoyan lo que Laurel encontró de manera intuitiva a partir de los histogramas? ¿Qué medida de tendencia central parece más apropiada en esta situación? Calcule las ventas totales de la compañía en dólares para los últimos cuatro trimestres. ¿Tiene razón Stan al afirmar que las ventas totales están bien?
2. Calcule la media del número de pedidos diarios y el tamaño de los pedidos para el centro de distribución 3 (Pennsylvania) durante los últimos cuatro trimestres. ¿Este centro de distribución muestra tendencias parecidas a las de toda la compañía? ¿Los planes de Laurel de hacer una investigación del desempeño de cada centro de distribución son una buena idea?

El jueves en la tarde, Laurel encontró a Hal en su oficina y le dio una breve descripción de sus hallazgos.

"Todo esto es muy interesante", respondió Hal. "Me gustaría saber qué opina el resto de la directiva en la reunión del lunes. ¿Crees poder montar una pequeña presentación? Tendría que ser muy clara en cuanto a las conclusiones y no quedarse mucho en las estadísticas."

"Seguro", acordó Laurel. "Todavía quiero hacer algunas pruebas de variabilidad; luego estaré en condiciones de armar todo el rompecabezas. Nos vemos el lunes."

3. Determine los rangos intercuartiles del tamaño promedio de los pedidos en cada trimestre. Compárelos con el rango total en cada caso.
4. Utilizando los datos sin procesar, calcule la varianza y la desviación estándar de la muestra por trimestre, para el número de pedidos y el tamaño promedio de esos pedidos. Calcule el coeficiente de variación para cada trimestre.
5. a) Utilice el teorema de Chebyshev para determinar el rango del número diario de pedidos y del tamaño promedio de éstos para el segundo trimestre de 1992 que incluirá al menos 75% de los datos.

   b) Examine los histogramas representados gráficamente en el ejercicio correspondiente del capítulo 2 y compárelos con los rangos de Chebyshev calculados en el inciso a). ¿Qué tan preciso es el teorema de Chebyshev para establecer el rango en cada caso?
6. Considerando cada centro de distribución por separado, calcule el coeficiente de variación para el número de pedidos y su tamaño promedio en el periodo completo de 12 meses. ¿Existen diferencias significativas entre las dispersiones relativas experimentadas por cada centro de distribución?
7. ¿De qué manera presentaría sus resultados a la junta directiva? ¿Qué recomendaciones podría hacer con respecto a las promociones, a las futuras recolecciones de datos, etcétera?

Repaso del capítulo

● Términos introducidos en el capítulo 3

Análisis exploratorio de datos (AED) Métodos para analizar datos que requieren de muy pocas suposiciones principales.

Clase de la mediana Clase de una distribución de frecuencias que contiene el valor mediano de un conjunto de datos.

Codificación Método para calcular la media de datos agrupados mediante la recodificación de los valores de los puntos medios de las clases en valores más sencillos.

Coeficiente de variación Medida relativa de la dispersión, que puede compararse para diferentes distribuciones y que expresa la desviación estándar como porcentaje de la media.

Cuartiles Fractiles que dividen los datos en cuatro partes iguales.

Curtosis Medida de lo puntiagudo de una distribución de puntos.

Deciles Fractiles que dividen los datos en 10 partes iguales.

Desviación estándar Raíz cuadrada positiva de la varianza; medida de dispersión con las mismas unidades que los datos originales, más que en las unidades al cuadrado en que se expresa la varianza.

Dispersión La separación o variabilidad de un conjunto de datos.

Distribución bimodal Distribución de datos puntuales en la que dos valores ocurren con más frecuencia que los demás valores del conjunto de datos.

Estadísticos Medidas numéricas que describen las características de una muestra. Representados por caracteres latinos.

Estadística sumaria Números sencillos que describen ciertas características de un conjunto de datos.

Fractil En una distribución de frecuencias, es la posición de un valor en, o más grande que, una fracción dada de los datos.

Gráfica de caja Técnica gráfica de AED empleada para resaltar el centro y los extremos de un conjunto de datos.

Media Medida de tendencia central que representa el promedio aritmético de un conjunto de observaciones.

Media geométrica Medida de tendencia central utilizada para medir la tasa promedio de cambio o de crecimiento de alguna cantidad, se calcula tomando la *n*-ésima raíz del producto de *n* valores que representan el cambio.

Media ponderada Promedio que se calcula con el fin de tomar en cuenta la importancia de cada valor con respecto al total, esto es, un promedio en el que cada valor de observación se pondera con algún índice de su importancia.

Mediana Punto situado a la mitad del conjunto de datos, medida de localización que divide al conjunto de datos en dos partes iguales.

Medida de dispersión Medida que describe cómo se dispersan o separan las observaciones de un conjunto de datos.

Medida de distancia Medida de dispersión en términos de la diferencia entre dos valores del conjunto de datos.

Medida de tendencia central Medida que indica el valor que debe esperarse para un dato típico o situado en el centro.

Moda El valor que ocurre más a menudo un conjunto de datos. Está representado por el punto más alto de la curva de la distribución de un conjunto de datos.

Parámetros Valores numéricos que describen las características de una población completa; suelen representarse con letras griegas.

Percentiles Fractiles que dividen los datos en 100 partes iguales.

Rango Distancia entre los valores más bajo y más alto de un conjunto de datos.

Rango intercuartil Diferencia entre los valores del primer y tercer cuartiles; esta diferencia representa el rango de la mitad central del conjunto de datos.

Rango interfractil Medida de la dispersión entre dos fractiles de una distribución; es decir, la diferencia entre los valores de dos fractiles.

Resultado estándar Expresión de una observación en términos de unidades de desviación estándar arriba o abajo de la media; es decir, la transformación de una observación al restarle la media y dividirla entre la desviación estándar.

Sesgo Grado en que una distribución de puntos está concentrada en un extremo o en el otro; falta de simetría.

Simétrica Característica de una distribución en la que cada mitad es la imagen de espejo de la otra.

Teorema de Chebyshev No importa qué forma tenga la distribución, al menos 75% de los valores de la población caerán dentro de dos desviaciones estándar a partir de la media, y al menos 89% caerá dentro de tres desviaciones estándar.

Varianza Medida del cuadrado de la distancia promedio entre la media y cada observación de la población.

● Ecuaciones introducidas en el capítulo 3

■ **3-1**
$$\mu = \frac{\Sigma x}{N}$$

La *media aritmética de la población* es igual a la suma de los valores de todos los elementos de la población (Σx) dividida entre el número total de elementos que componen la población (N).

■ **3-2**
$$\bar{x} = \frac{\Sigma x}{n}$$

Para calcular la *media aritmética de la muestra*, sume los valores de todos los elementos de la muestra (Σx) y divida el resultado entre el número total de elementos de la muestra (n).

■ **3-3**
$$\bar{x} = \frac{\Sigma(f \times x)}{n}$$

Para encontrar la *media aritmética de la muestra con datos agrupados*, calcule los puntos medios (x) de cada clase de la muestra. Luego multiplique cada punto medio por la frecuencia (f) de observaciones de cada clase, sume (Σ) todos estos productos y divida la suma entre el número total de observaciones de la muestra (n).

■ **3-4**
$$\bar{x} = x_0 + w\,\frac{\Sigma(u \times f)}{n}$$

Esta fórmula nos permite calcular la *media aritmética de la muestra de datos agrupados* mediante el uso de códigos, con el fin de evitarnos trabajar con puntos medios muy grandes o inconvenientes. Asigne estos códigos (u) de la manera siguiente: fije el valor cero al punto medio (denotado por x_0), enteros positivos consecutivos a los puntos medios mayores a x_0 y enteros negativos consecutivos a los puntos medios menores. Luego multiplique el código asignado a cada clase (u) por la frecuencia (f) de las observaciones de cada clase y sume (Σ) todos los productos. Divida el resultado entre el número total de observaciones de la muestra (n), multiplique por el ancho numérico del intervalo de clase (w) y sume el valor del punto medio correspondiente al código cero (x_0).

■ **3-5**
$$\bar{x}_w = \frac{\Sigma(w \times x)}{\Sigma w}$$

La *media ponderada*, $\bar{x}_n$, es un promedio que toma en cuenta qué tan importante es cada valor respecto al total. Podemos calcular este promedio multiplicando el peso, o ponderación, de cada elemento (w) por el elemento correspondiente (x), sumando el resultado de todos esos productos (Σ) y dividiendo esta cantidad entre la suma de todas las ponderaciones (Σw).

■ **3-6**
$$\text{M.G.} = \sqrt[n]{\text{producto de todos los valores } x}$$

La *media geométrica* o M.G. es adecuada siempre que necesitemos medir la tasa promedio de cambio (tasa de crecimiento) en un periodo. En esta ecuación, n es igual al número de valores x que aparecen en el problema.

■ **3-7**
$$\text{Mediana} = \left(\frac{n+1}{2}\right)\text{-ésimo término del arreglo de datos}$$

donde n = número de elementos del ordenamiento de datos

La *mediana* es un solo valor que mide el elemento central del conjunto de datos. La mitad de las observaciones quedan arriba de la mediana y la otra mitad abajo. Si el conjunto de datos contiene un número impar de observaciones, el elemento de en medio es la mediana. Para un número par de elementos, la mediana es el promedio de las dos observaciones de en medio. Utilice esta ecuación cuando los datos no están agrupados.

■ **3-8**
$$\tilde{m} = \left(\frac{(n+1)/2 - (F+1)}{f_m}\right)w + L_m$$

Esta fórmula nos permite encontrar la *mediana de la muestra de datos agrupados*. En ella, n es igual al número total de observaciones de la distribución; F es la suma de todas las frecuencias de clase hasta la clase mediana, sin incluirla; f_m es la frecuencia de las observaciones de la clase de la mediana; w es el ancho de intervalos de clase, y L_m es el límite inferior del intervalo de la clase de la mediana.

■ **3-9**
$$Mo = L_{Mo} + \left(\frac{d_1}{d_1+d_2}\right)w$$

La *moda* es el valor que con se repite más frecuencia en el conjunto de datos. Para hallar la *moda de datos agrupados* (denotada con *Mo*), utilice esta fórmula y tome a L_{Mo} igual al límite inferior de la clase modal; d_1 como la frecuencia de la clase modal menos la frecuencia de la clase que está inmediatamente abajo de ella; d_2 igual a la frecuencia de la clase modal menos la frecuencia de la clase que está inmediatamente arriba de ella, y w como el ancho del intervalo de la clase modal.

■ **3-10**
$$\text{Rango} = \begin{array}{c}\text{valor de la}\\ \text{observación más alta}\end{array} - \begin{array}{c}\text{valor de la}\\ \text{observación más baja}\end{array}$$

El *rango* es la diferencia entre los valores más alto y más bajo de una distribución de frecuencias.

■ **3-11**
$$\text{Rango intercuartil} = Q_3 - Q_1$$

El *rango intercuartil* mide aproximadamente qué tan lejos de la mediana debemos desplazarnos a ambos lados antes de poder incluir una mitad de los valores del conjunto de datos. Para calcular este rango, divida los datos en cuatro partes iguales. Los *cuartiles* (Q) son los valores más altos de cada una de esas cuatro partes. El *rango intercuartil* es la diferencia entre los valores del primer y el tercer cuartil (Q_1 y Q_3).

■ **3-12**
$$\sigma^2 = \frac{\Sigma(x-\mu)^2}{N} = \frac{\Sigma x^2}{N} - \mu^2$$

Esta fórmula nos permite calcular la *varianza de la población*, una medida del *cuadrado* de la distancia promedio entre la media y cada observación de la población. La expresión de en medio, $\frac{\Sigma(x-\mu)^2}{N}$ es la definición de σ^2. La última expresión, $\frac{\Sigma x^2}{N-\mu^2}$ es matemáticamente equivalente a la definición, pero a menudo es mucho más conveniente usarla, debido a que nos libera del cálculo de las desviaciones de la media.

■ **3-13**
$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\Sigma(x-\mu)^2}{N}} = \sqrt{\frac{\Sigma x^2}{N} - \mu^2}$$

La desviación estándar de la población, σ, es la raíz cuadrada de la varianza de la población. Es un parámetro más útil que la varianza, debido a que se expresa en las mismas unidades que los datos (mientras que las unidades de la varianza son el cuadrado de las unidades de los datos). La desviación estándar es siempre la raíz cuadrada *positiva* de la varianza.

■ **3-14**
$$\text{Resultado estándar de la población} = \frac{x-\mu}{\sigma}$$

El *resultado estándar* de una observación es el número de desviaciones estándar que la observación se separa hacia abajo o hacia arriba de la media de la distribución. El resultado estándar nos permite hacer comparaciones entre los elementos de distribuciones que difieren en orden de magnitud o en las unidades empleadas. Utilice la ecuación 3-14 para encontrar el resultado estándar de una observación de una *población*.

■ **3-15**
$$\sigma^2 = \frac{\Sigma f(x-\mu)^2}{N} = \frac{\Sigma f x^2}{N} - \mu^2$$

Esta fórmula, en cualquiera de sus formas, nos permite calcular la *varianza de datos ya agrupados* en una distribución de frecuencias. En ella, f representa la frecuencia de la clase y x es el punto medio.

■ **3-16**
$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\Sigma f(x-\mu)^2}{N}} = \sqrt{\frac{\Sigma f x^2}{N} - \mu^2}$$

Tome la raíz cuadrada de la varianza y obtendrá la *desviación estándar de datos agrupados*.

■ **3-17**
$$s^2 = \frac{\Sigma(x - \bar{x})^2}{n - 1} = \frac{\Sigma x^2}{n - 1} - \frac{n\bar{x}^2}{n - 1}$$

Para calcular la *varianza de la muestra*, utilice la misma fórmula de la ecuación 3-12, sustituyendo μ con $\bar{x}$ y N con $n - 1$. En el capítulo 7 se explica por qué utilizamos $n - 1$ en lugar de n, para calcular la varianza de la muestra.

■ **3-18**
$$s = \sqrt{s^2} = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n - 1}} = \sqrt{\frac{\Sigma x^2}{n - 1} - \frac{n\bar{x}^2}{n - 1}}$$

La *desviación estándar de la muestra* es la raíz cuadrada de la varianza de la muestra. Es parecida a la ecuación 3-13, excepto que μ se sustituye por la media de la muestra $\bar{x}$ y N se cambia por $n - 1$.

■ **3-19**
$$\text{Resultado estándar de la muestra} = \frac{x - \bar{x}}{s}$$

Utilice esta ecuación para encontrar el resultado estándar de una observación en una *muestra*.

■ **3-20**
$$\text{Coeficiente de variación de la población} = \frac{\sigma}{\mu}(100)$$

El *coeficiente de variación* es una medida relativa de dispersión que nos permite comparar dos distribuciones. Relaciona la desviación estándar y la media mediante la expresión de la desviación estándar como porcentaje de la media.

● Ejercicios de repaso

■ **3-83** El departamento de pesos y medidas de la oficina de agricultura de un estado midió la cantidad de granola que se vende en paquetes de 4 onzas y registró los siguientes datos:

4.01	4.00	4.02	4.02	4.03	4.00	3.98	3.99	3.99	4.01
3.99	3.98	3.97	4.00	4.02	4.01	4.02	4.00	4.01	3.99

Si la muestra es representativa de todos los tipos de granola que vende este fabricante, ¿cuál es el rango de los pesos para 95% de los paquetes?

■ **3-84** ¿Qué reacción tendría usted si un aficionado al fútbol americano le dijera lo siguiente? "Los Raiders de Rockland tienen un promedio de 3.6 yardas de recorrido por tierra. Como sólo necesitan 10 yardas para anotar y tienen cuatro oportunidades para lograrlo, la anotación es segura, siempre y cuando mantengan su forma de jugar por tierra."

■ **3-85** ¿Qué respondería al siguiente comentario?: "La variabilidad no es un factor importante, porque aun cuando el resultado es más incierto, de todos modos se tiene una misma posibilidad de caer arriba o abajo de la mediana. En consecuencia, en promedio, el resultado será el mismo."

■ **3-86** A continuación se presentan tres partes del presupuesto de una año para la defensa; a cada una de ellas, el Congreso estadounidense asignó la misma cantidad de financiamiento:

a) Salario de los funcionarios (total).
b) Mantenimiento de la flota aérea.
c) Adquisiciones de alimentos (total).

Tomando en cuenta la distribución de los resultados posibles para los gastos reales en cada una de estas áreas, haga corresponder cada sección a una de las curvas de la figura 3-9. Fundamente su respuesta.

■ **3-87** La compañía Ed's Sports Equipment tiene en existencia dos categorías de sedal de pesca. Los datos sobre cada categoría son los siguientes:

	Resistencia media de prueba (kg)	Desviación estándar
Master	40	Valor exacto desconocido, pero se estima que es muy grande
Super	30	Valor exacto desconocido, pero se estima que es muy pequeño

Si usted se dispone a pescar un tipo de pez cuyo peso promedio ha sido 25 kg en esta temporada, ¿con cuál de los dos sedales tiene más posibilidad de atrapar una cantidad mayor de peces?

■ **3-88** El vicepresidente de ventas de Vanguard Products ha estado estudiando los registros correspondientes al desempeño de sus representantes de ventas. Se ha dado cuenta que en los últimos dos años, el nivel promedio de ventas por representante ha permanecido igual, mientras que la distribución de los niveles de ventas se ha ampliado. Los niveles de ventas de los agentes de la compañía para ese mismo periodo tienen variaciones significativamente más grandes respecto a la media que en cualquier otro periodo de dos años anterior al estudiado. ¿A qué conclusiones puede llegar basándose en esas observaciones?

■ **3-89** Los automóviles nuevos vendidos en diciembre por ocho distribuidores de Ford situados en un radio de 80 kilómetros de Canton, Ohio, pueden describirse en el siguiente conjunto de datos:

200 156 231 222 96 289 126 308

a) Calcule el rango, el rango intercuartil y la desviación estándar de estos datos.
b) ¿Cuál de las tres medidas que calculó para responder al inciso a) describe mejor la variabilidad de los datos?

■ **3-90** Dos economistas estudian las fluctuaciones del precio del oro. Uno examina el periodo correspondiente a 1968-1972; el otro estudia el correspondiente a 1975-1979. ¿Qué diferencias esperaría encontrar en la variabilidad de sus datos?

■ **3-91** La fábrica de botas para esquiar Downhill opera dos líneas de ensamble en sus plantas. El gerente de producción está interesado en mejorar la consistencia de la línea que posee la mayor variación. La línea número 1 produce un promedio mensual de 11,350 unidades, con una desviación estándar de 1,050. La 2 produce un promedio mensual de 9,935 unidades, con desviación estándar de 1,010. ¿Cuál de las dos líneas posee la mayor dispersión relativa?

■ **3-92** E1 30 de junio de 1992, la capitalización de nueve mercados de valores del Pacífico y Asia fue:

País	Capitalización (en miles de millones de dólares)
Filipinas	17
Indonesia	21
Tailandia	44
Singapur	50
Malasia	79
Corea del Sur	86
Taiwan	140
Hong Kong	178
Australia	203

Fuente: "Asian/Pacific Stock Markets", en The Chicago Tribune *(14 de diciembre de 1992), sec. 4, p. 3.*

a) Encuentre la media aritmética de los datos.
b) Encuentre la mediana de los datos.
c) Encuentre la moda de los datos.
d) ¿Cuál es la mejor medida de la tendencia central del conjunto de datos?
e) Encuentre la desviación estándar de los datos. (La población completa está incluida en ellos.)

■ **3-93** La estación Fish and Game en el lago Wylie tiene registros de los peces atrapados en el lago e informa sus hallazgos al Servicio Nacional de Pesca Deportiva. La pesca en libras de los 20 últimos días fue:

101 132 145 144 130 88 156 188 169 130
90 140 130 139 99 100 208 192 165 216

Calcule el rango, la varianza y la desviación estándar para estos datos. En este caso, ¿es el rango una buena medida de la variabilidad? ¿Por qué?

■ **3-94** El dueño de Records Anonymous, un comerciante de discos al menudeo, emplea dos fórmulas diferentes para pronosticar sus ventas mensuales. La primera fórmula tiene una falla promedio de 700 discos, con desviación estándar de 35. La segunda, de 300 discos, con desviación estándar de 16. ¿Cuál fórmula es relativamente menos precisa?

■ **3-95** Utilizando los siguientes datos de población, calcule el rango intercuartil, la varianza y la desviación estándar. ¿Qué concluye de sus respuestas acerca del comportamiento del costo de combustible para calentar?

Costo promedio por galón de combustible para calentar en ocho estados de la Unión Americana

1.89	1.66	1.77	1.83	1.71	1.68	1.69	1.73

■ **3-96** La siguiente tabla presenta la cantidad promedio de policías hombres y mujeres del Departamento de Policía de Nueva York que estuvieron en servicio cada día entre las 20:00 y las 24:00 horas en el barrio de Manhattan:

Lunes	2,950	Miércoles	2,900	Viernes	3,285	Domingo	2,975
Martes	2,900	Jueves	2,980	Sábado	3,430		

a) ¿Serían la varianza o la desviación estándar una buena medida de la variabilidad de estos datos?
b) ¿Qué encontró en el patrón del número de policías que ocasionó su respuesta al inciso a)?

■ **3-97** Un sicólogo escribió un programa de computación para simular la forma en que una persona responde a una prueba típica de CI (cociente de inteligencia). Para probar el programa, le dio a la computadora 15 formas distintas de una prueba de CI y calculó su coeficiente de inteligencia obtenido en cada forma.

Valores de CI

134	136	137	138	138
143	144	144	145	146
146	146	147	148	153

a) Calcule la media y la desviación estándar de los resultados de CI.
b) De acuerdo con el teorema de Chebyshev, ¿cuántos valores deberían estar entre 132.44 y 153.56? ¿Cuántos valores se encuentran realmente en ese intervalo?

■ **3-98** Liquid Concrete entrega mezcla de concreto lista en 40 camiones. El número de yardas cúbicas entregadas por cada camión cierto día fue el siguiente:

Yardas cúbicas

11.9	12.8	14.6	15.8	13.7	9.9	18.8	16.9	10.4	9.1
17.1	13.0	18.6	16.0	13.9	14.7	17.7	12.1	18.0	17.8
19.0	13.3	12.4	9.3	14.2	15.0	19.3	10.6	11.2	9.6
13.6	14.5	19.6	16.6	12.7	15.3	10.9	18.3	17.4	16.3

Enumere los valore en cada decil; 8% de los camiones entregaron menos de _____ yardas cúbicas.

■ **3-99** La asistencia a los 10 últimos partidos en casa de las Águilas de Baltimore fue la siguiente:

20,100	24,500	31,600	28,400	49,500
19,350	25,600	30,600	11,300	28,560

a) Calcule el rango, la varianza y la desviación estándar para estos datos.
b) ¿Algunos de los resultados que obtuvo en el inciso a) son un reflejo preciso de la variabilidad de los datos de asistencia a los partidos?
c) ¿Qué otra medida de variabilidad podría ser mejor?
d) Calcule el valor de la medida que sugirió en el inciso c).

■ **3-100** Matthews, Young y Asociados, una agencia de consultorías de Chapell Hill, tiene los siguientes registros que indican el número de días que cada uno de sus ocho consultores de planta cobró el último año:

212	220	230	210	228	229	231	219	221	222

a) Sin calcular el valor de las medidas, ¿cuál de ellas cree usted que le daría una mayor información acerca de esta distribución: el rango o la desviación estándar?
b) Tomando en cuenta la dificultad y el tiempo para calcular cada medida que revisó en el inciso a), ¿cuál sugeriría como la mejor?
c) ¿Qué haría que usted cambiara su opinión al respecto?

■ **3-101** La siguiente distribución de frecuencias resume los cambios de precios ocurridos el 24 de mayo de 1993 en todas las compañías que participaron en la Bolsa de Valores de Nueva York y cuyos nombres comienzan con *L* o con *R*.

Cambio de precio	Número de compañías con *L*	Número de compañías con *R*
−1.25 a −1.01	1	1
−1.00 a −0.76	1	1
−0.75 a −0.51	1	0
−0.50 a −0.26	7	5
−0.25 a −0.01	19	20
0.00	14	20
0.01 a 0.25	21	14
0.26 a 0.50	5	8
0.51 a 0.75	3	1
0.76 a 1.00	2	4
1.01 a 1.25	1	0

Fuente: The Wall Street Journal *(25 de mayo de 1993): C4-C5.*

a) Encuentre la media aritmética de las dos distribuciones.
b) Encuentre su mediana.
c) Encuentre su moda.
d) ¿Cuál es la mejor medida de tendencia central para cada distribución?
e) Encuentre la desviación estándar de las dos distribuciones (cada grupo es una población completa).
f) Utilice sus coeficientes de variación para determinar qué distribución tiene menor variabilidad relativa.

■ **3-102** Larsen Equipment Rental proporciona a los contratistas las herramientas que necesitan sólo por unos días, como sierras para concreto. Cuando el equipo se descompone al estar rentado, debe considerarse fuera de servicio hasta que se repara. Con frecuencia se hace rápido, pero algunas veces tarda mientras llegan las refacciones. Es útil hacer un análisis del tiempo perdido para planear el inventario. Los registros de descomposturas en el último año son:

Grupo de equipos	Días descompuesto	Grupo de equipos	Días descompuesto
1	2	8	8
2	19	9	29
3	14	10	6
4	21	11	0
5	5	12	4
6	7	13	4
7	11	14	10

a) ¿Cuál fue el tiempo medio de descomposturas el año pasado para los grupos de equipos?
b) ¿Cuál fue la mediana?

■ **3-103** Larsen (vea el ejercicio 3-102) acaba de obtener la siguiente información adicional:

Grupo de equipos	Piezas de maquinaria	Grupo de equipos	Piezas de maquinaria
1	1	8	5
2	3	9	8
3	1	10	2
4	4	11	2
5	2	12	6
6	1	13	1
7	1	14	1

a) ¿Cuál es el tiempo promedio de descompostura por pieza de maquinaria?
b) ¿Cuál es el tiempo promedio de descompostura por pieza de maquinaria para cada grupo al clasificarlas por grupo?
c) ¿Cuánto grupos tienen un tiempo de descomposturas arriba del promedio por pieza de maquinaria?

■ **3-104** Compare y contraste, en términos generales, la posición central y el sesgo de las distribuciones del nivel de lectura, en número de lectores, por volumen para todas las:
a) Revistas mensuales.
b) Revistas noticiosas semanales.
c) Revistas médicas mensuales, distribuidas a nivel nacional.

■ **3-105** Compare y contraste, en términos generales, la tendencia central y el sesgo de las distribuciones concernientes a la cantidad de impuestos (en dólares) pagados por todos

a) los individuos que solicitan reembolso federal en Estados Unidos, donde la categoría superior de impuestos es del 28%.
b) los individuos que pagan impuestos estatales en Carolina del Norte, donde la categoría superior de impuestos es 7%.
c) los individuos que pagan impuestos por derecho de aeropuerto (incluido en el precio del boleto de avión) en el Aeropuerto Internacional JFK de la ciudad de Nueva York.

■ **3-106** Allison Barrett realiza análisis estadísticos para un equipo de carreras automovilísticas. A continuación, presentamos las cifras en millas por galón del gasto de combustible de sus automóviles en carreras recientes:

4.77	6.11	6.11	5.05	5.99	4.91	5.27	6.01
5.75	4.89	6.05	5.22	6.02	5.24	6.11	5.02

a) Calcule la mediana del consumo de combustible.
b) Calcule la media del mismo consumo.
c) Agrupe los datos en cinco clases de igual tamaño. ¿Cuál es el valor del consumo de combustible para la clase modal?
d) ¿Cuál de las tres medidas de tendencia central es la que mejor puede servirle a Allison cuando haga un pedido de combustible? Explique su respuesta.

■ **3-107** Pidieron a Clara Chávez, analista del Servicio Interno de Contribuciones estadounidense, que describa al contribuyente "promedio" de la Unión Americana en términos del ingreso bruto anual. Clara tiene datos sumarios en los que el contribuyente está clasificado en diferentes clases según el impuesto que paga. ¿Cuál medida de tendencia central deberá utilizar?

■ **3-108** La comercializadora de flores Emmot Bulb Co., vende bolsas sorpresa con bulbos de flores. Las bolsas se venden según su peso; en consecuencia, el número de bulbos en cada bolsa puede variar, dependiendo de las variedades incluidas. El número de bulbos que hay en cada bolsa de una muestra de 20 son:

21	33	37	56	47
36	23	26	33	37
25	33	32	47	34
26	37	37	43	45

a) ¿Cuáles son la media y la mediana del número de bulbos por bolsa?
b) Con base en su respuesta, ¿qué puede concluir acerca de la forma de la distribución del número de bulbos por bolsa?

■ **3-109** Un ingeniero probó nueve muestras de cada uno de tres diseños de soporte para un nuevo torno eléctrico. Los datos siguientes corresponden al número de horas que tardó cada soporte en fallar con el motor del torno funcionando continuamente a su máxima potencia, con una carga equivalente a 1.9 veces su capacidad diseñada.

Diseño		
A	**B**	**C**
16	18	31
16	27	16
53	23	42
15	21	20
31	22	18
17	26	17
14	39	16
30	17	15
20	28	19

a) Calcule la media y la mediana para cada grupo.
b) Con base en su respuesta, ¿cuál diseño es el mejor y por qué?

■ **3-110** La Table Spice Co. está instalando una criba en una sección de su nueva planta de procesamiento para separar hojas, tierra e insectos de una clase costosa de semilla de especia, que recibe a granel de

los cultivadores. La empresa puede utilizar una malla gruesa de 3.5 mm o una más fina de 3 mm. La malla más fina retendrá más materia inútil, pero también eliminará más semillas. La malla más abierta dejará pasar más basura y eliminará menos semillas. La compañía tiene la siguiente información tomada de una muestra de los desechos:

Tamaño de desechos (en milímetros)	Frecuencia
1.0 menor que	12
1.01-1.5	129
1.51-2.0	186
2.01-2.5	275
2.51-3.0	341
3.01-3.5	422
3.51-4.0	6,287
4.01-4.5	8,163
4.51-5.0	6,212
5.01-5.5	2,416
mayor que 5.5	1,019

a) ¿Cuál es el tamaño de la mediana de basura y el tamaño de la clase modal?
b) ¿Cuál malla usaría usted, según el inciso a), si desea eliminar al menos la mitad de la basura?

■ **3-111** WordPerfect, el desarrollador del popular software de procesamiento de texto, logró una gran participación de mercado mediante una estrategia de negocios que incluyó soporte técnico telefónico gratis ilimitado. En 1993, la compañía reportó el siguiente resumen estadístico:

Promedio diario de llamadas a soporte	16,500
Tiempo promedio "en espera"	1 minuto 36 segundos
Total de llamadas contestadas	3,976,951

Fuente: WordPerfect Report, *primavera de 1993.*

Suponga que contestan llamadas los 365 días del año, ¿el "promedio" reportado es una media aritmética? Suponga que un representante técnico típico puede atender 165 llamadas por turno, ¿tiene sentido programar 100 personas por día?

■ **3-112** Las siguientes son las cantidades promedio (en dólares) que gasta en comida cada línea aérea por pasajero:

American	7.41
United	7.24
Northwest	5.15
TWA	5.09
Delta	4.61
Continental	2.77
USAir	2.68
American West	2.00
Southwest	0.14

Fuente: "The Going Rate", The Wall Street Journal *(23 de junio de 1995). B11.*

¿Cuál es la media y la mediana del costo por pasajero? ¿Cuál sería la mejor cifra para usar en una línea aérea que desarrolla su plan de negocios?

■ **3-113** Merrill Lynch, la empresa más importante de corredores de bolsa, encargó un estudio de la riqueza de las familias estadounidenses. A partir de los datos del censo de Estados Unidos para más de 38,000 familias, el doctor Joseph Anderson concluyó que el nivel de la mediana de los bienes financieros (fondos para inversión, excluyendo bienes raíces) era $1,000, y la media aproximada era $30,000. Como $1,000 es menos que la cantidad mínima necesaria para invertir en valores, ¿deben los administradores de la empresa pensar en dejar de vender acciones al público en general?

Fuente: B. Wysocki, "Many Baby Boomers Save Little, May Run Into Trouble Later On", *The Wall Street Journal* (5 de junio de 1995): 1. Los detalles se tomaron de un comunicado de prensa de Merrill Lynch, "New Data Shows Wealth of American Families at 'Woeful Low' " (21 de diciembre de 1994).

capítulo 7

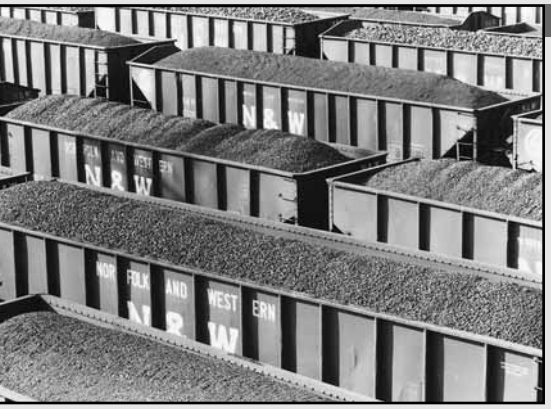
ESTIMACIÓN

Objetivos

- Aprender cómo hacer estimaciones de ciertas características de una población a partir de muestras
- Aprender las fortalezas y limitaciones de las estimaciones puntuales y las estimaciones de intervalo
- Calcular qué tan precisas son en realidad nuestras estimaciones
- Aprender a utilizar la distribución *t* para hacer estimaciones de intervalo en algunos casos en los que la distribución normal no se puede utilizar
- Calcular el tamaño de muestra requerido para cualquier nivel deseado de precisión en la estimación

Contenido del capítulo

7.1 Introducción 274
7.2 Estimaciones puntuales 277
7.3 Estimaciones de intervalo: conceptos básicos 281
7.4 Estimaciones de intervalo e intervalos de confianza 285
7.5 Cálculo de estimaciones de intervalo de la media a partir de muestras grandes 288
7.6 Cálculo de estimaciones de intervalo de la proporción a partir de muestras grandes 293
7.7 Estimaciones de intervalos con la distribución *t* 297
7.8 Determinación del tamaño de muestra para la estimación 303
- Estadística en el trabajo 309
- Ejercicio de base de datos computacional 309
- Del libro de texto al mundo real 311
- Términos introducidos en el capítulo 7 312
- Ecuaciones introducidas en el capítulo 7 313
- Ejercicios de repaso 313



Como parte del proceso de asignar el presupuesto del año siguiente, el administrador de la planta generadora de energía eléctrica Far Point debe estimar la cantidad de carbón que requerirá para este año. El año anterior, la planta casi se quedó sin combustible, de modo que el administrador está reticente a solicitar el mismo presupuesto de nuevo. Sin embargo, el administrador de la planta siente que el uso de los datos registrados le ayudará para *estimar* el número de toneladas de carbón que debe pedir. Una muestra aleatoria de 10 semanas de operación de la planta seleccionadas de los últimos cinco años produjo un consumo medio de 11,400 toneladas semanales, con una desviación estándar de la muestra de 700 toneladas por semana. Con los datos que tiene a su disposición y los métodos que se estudian en este capítulo, el administrador de la planta puede hacer una buena estimación de la cantidad que debe pedir este año, e incluso tener una idea de qué tan precisa es la estimación. ■

7.1 Introducción

Todo el mundo hace estimaciones. Cuando está por cruzar una calle, hace una estimación de la velocidad del automóvil que se acerca, de la distancia que hay entre usted y el auto y de su propia velocidad. Habiendo hecho rápidamente todas estas estimaciones, usted decide si espera, camina o corre.

Razones para hacer estimaciones

Los administradores también deben hacer estimaciones rápidas. El resultado de estas estimaciones puede afectar sus organizaciones de manera tan seria como el resultado de su decisión de cruzar la calle. Los jefes de departamento de una universidad hacen estimaciones acerca de las inscripciones para el semestre siguiente en las materias. Los directores de crédito estiman si un cliente pagará o no sus débitos. Los futuros compradores de casa hacen estimaciones concernientes al comportamiento de las tasas de interés de los préstamos hipotecarios. Todas estas personas hacen estimaciones sin preocuparse de si son científicas o no, pero con la esperanza de que las estimaciones tengan una semejanza razonable con el resultado.

Los administradores utilizan estimaciones porque, hasta en los asuntos más triviales, deben tomar decisiones racionales sin contar con la información pertinente completa y con una gran incertidumbre de lo que el futuro pueda deparar. Como ciudadanos instruidos y profesionales, podremos hacer estimaciones más útiles si aplicamos las técnicas descritas en este capítulo y los que le siguen.

Elaboración de inferencias estadísticas

El material sobre teoría de probabilidad que se presentó en los capítulos 4, 5 y 6 constituye la base de la *inferencia estadística*, rama de la estadística que se ocupa del uso de los conceptos de probabilidad para manejar la incertidumbre en la toma de decisiones. La inferencia estadística está basada en la *estimación*, concepto que se introduce en este capítulo, y en las *pruebas de hipótesis*, que es el tema de los capítulos 8, 9 y 10. Tanto en la estimación como en las pruebas de hipótesis, haremos inferencias acerca de las características de las poblaciones a partir de la información proporcionada por las muestras.

Uso de muestras

¿De qué manera los administradores utilizan estadísticas para estimar los parámetros de una población? El jefe de departamento de alguna universidad intenta estimar el número de inscripciones que tendrá el siguiente semestre a partir de las inscripciones actuales en los mismos cursos. El director de un departamento de crédito intentará estimar el valor crediticio de los futuros clientes a partir de una muestra de sus hábitos de pago. El comprador de una casa intenta estimar el curso futuro de las tasas de interés mediante la observación de su comportamiento actual. En cada caso, alguien trata de inferir algo acerca de una población a partir de la información adquirida de una muestra.

Estimación de parámetros

En este capítulo introducimos métodos que nos permiten estimar con precisión razonable la *proporción de la población* (la fracción de la población que posee una característica dada) y la *media de la población*. Calcular la proporción exacta o la media exacta sería una meta imposible. Pero, a pesar de ello, seremos capaces de hacer una estimación, establecer una afirmación respecto al error que tal vez acompañará a esta estimación, y poner en marcha algunos controles para evitar dicho error en la medida de lo posible. Como tomadores de decisiones, nos veremos forzados, en ocasiones, a confiar en nuestros presentimientos. Sin embargo, en otras situaciones, en las que dispongamos de información y podamos aplicar los conceptos de estadística, tendremos mejores resultados.

Tipos de estimaciones

Definición de *estimación puntual*

Podemos hacer dos tipos de estimaciones concernientes a una población: una estimación *puntual* y una estimación de *intervalo*. **Una estimación puntual es un solo número que se utiliza para estimar un parámetro de población desconocido.** Si, mientras observa al primer integrante de un equipo de fútbol americano salir al campo de juego, se dice: "¡Caramba! Apuesto a que el peso promedio de los jugadores defensivos es de 125 kilogramos", usted ha hecho una estimación puntual. El jefe de departamento de una universidad estaría haciendo una estimación puntual si afirmara: "Nuestros datos actuales indican que en esta materia tendremos 350 estudiantes el siguiente semestre."

Limitaciones de las estimaciones puntuales

A menudo, una estimación puntual es insuficiente debido a que sólo tienen dos opciones: es correcta o está equivocada. Si le dicen solamente que la afirmación sobre la inscripción está equivocada, no sabe *qué tanto* está mal y no puede tener la certeza de que la estimación es confiable. Si se entera de que sólo está errada por 10 estudiantes, podría aceptar a 350 estudiantes como una buena estimación de la inscripción futura. Pero si está equivocada en 90 estudiantes, la rechazaría como estimación de la inscripción futura. Entonces, una estimación puntual es mucho más útil si viene acompañada por una estimación del error que podría estar implicado.

Definición de *estimación de intervalo*

Una estimación de intervalo es un rango de valores que se utiliza para estimar un parámetro de la población. Una estimación de este tipo indica el error de dos maneras: por la extensión del intervalo y por la probabilidad de que el verdadero parámetro poblacional se encuentre dentro del intervalo. En este caso, el jefe de departamento diría algo como lo siguiente: "Estimo que la inscripción real de este curso para el próximo semestre estará entre 330 y 380, y es muy probable que la inscripción exacta caiga dentro de este intervalo." Con esto tiene una mejor idea de la confiabilidad de su estimación. Si el curso se imparte en grupos de 100 estudiantes cada uno y si, tentativamente, se han programado cinco cursos, entonces, de acuerdo con la estimación, puede cancelar uno de los grupos y abrir uno optativo.

Estimador y estimaciones

Definición de *estimador*

Cualquier estadístico de la muestra que se utilice para estimar un parámetro poblacional se conoce como *estimador*, es decir, **un estimador es un estadístico de la muestra utilizado para estimar un parámetro poblacional**. La media de la muestra $\bar{x}$ puede ser un estimador de la media de la población μ, y la proporción de la muestra se puede utilizar como un estimador de la proporción de la población. También es posible emplear el rango de la muestra como un estimador del rango de la población.

Definición de *estimación*

Cuando hemos observado un valor numérico específico de nuestro estimador, nos referimos a ese valor como una estimación. En otras palabras, **una estimación es un valor específico observado de un estadístico**. Hacemos una estimación si tomamos una muestra y calculamos el valor que toma nuestro estimador en esa muestra. Suponga que calculamos la lectura media de un odómetro (kilometraje) a partir de una muestra de taxis en servicio y encontramos que es 156,000 kilómetros. Si utilizamos este valor específico para estimar el kilometraje de la flotilla de taxis completa, el valor obtenido de 156,000 kilómetros sería una estimación. En la tabla 7-1 ilustramos varias poblaciones, parámetros, estimadores y estimaciones.

Tabla 7-1

Poblaciones, parámetros, estimadores y estimaciones

Población en la que estamos interesados	Parámetros de población que deseamos estimar	Estadístico de la muestra que utilizaremos como estimador	Estimación que realizamos
Empleados de una fábrica de muebles	Rotación media de empleados por año	Rotación media de empleados en un mes	8.9% de rotación por año
Candidatos a gerente la ciudad de Chapel Hill	Educación formal media (años)	Educación formal media de cada quinto solicitante	17.9 años de educación formal
Adolescentes de una comunidad dada	Proporción que tiene antecedentes penales	Proporción de una muestra de 50 adolescentes que tiene antecedentes penales	0.02, o 2%, tienen antecedentes penales

Criterios para seleccionar un buen estimador

Cualidades de un buen estimador

Algunos estadísticos son mejores estimadores que otros. Afortunadamente, podemos evaluar la calidad de un estadístico como estimador mediante el uso de cuatro criterios:

1. **Insesgado.** Ésta es una propiedad deseable para un buen estimador. El término *insesgado* se refiere al hecho de que una media de la muestra es un estimador no sesgado de una media de la población porque **la media de la distribución muestral de las medias de las muestras tomadas de la misma población es igual a la media de la población misma**. Podemos decir que un estadístico es un estimador insesgado (o no sesgado) si, en promedio, tiende a tomar valores que están arriba del parámetro de la población que se está estimando con la misma frecuencia y la misma extensión con la que tiende a asumir valores abajo del parámetro poblacional que se está estimando.
2. **Eficiencia.** Otra propiedad deseable de un buen estimador es que sea eficiente. La *eficiencia* se refiere al tamaño del error estándar del estadístico. Si comparamos dos estadísticos de una muestra del mismo tamaño y tratamos de decidir cuál de ellas es un estimador más eficiente, escogeríamos la estadística que tuviera el menor error estándar o la menor desviación estándar de la distribución muestral. Suponga que escogemos una muestra de un tamaño determinado y debemos decidir si utilizamos la media de la muestra o la mediana de la muestra para estimar la media de la población. Si calculamos el error estándar de la media de la muestra y encontramos que es 1.05, y luego calculamos el error estándar de la mediana de la muestra y tenemos que éste es 1.6, diríamos que la media de la muestra es un *estimador más eficiente* de la media poblacional *ya que su error estándar es menor.* Tiene sentido pensar que un estimador con un error estándar menor (con menos variación) tendrá mayor oportunidad de producir una estimación más cercana al parámetro poblacional que se está considerando.
3. **Consistencia.** Una estadística es un estimador consistente de un parámetro de población si *al aumentar el tamaño de la muestra, se tiene casi la certeza de que el valor de la estadística se aproxima bastante al valor del parámetro poblacional.* Si un estimador es consistente, se vuelve más confiable al tener tamaños de muestra más grandes. Si usted se pregunta acerca de la posibilidad de aumentar el tamaño de la muestra para obtener más información sobre un parámetro poblacional, averigüe primero si su estadístico es un estimador consistente o no. Si no lo es, desperdiciará tiempo y dinero al tomar muestras más grandes.
4. **Suficiencia.** Un estimador es *suficiente* si utiliza tanta información de la muestra que ningún otro estimador puede extraer información adicional acerca del parámetro de población que se está estimando.

Presentamos estos criterios con anticipación para que no pierda de vista el cuidado que los especialistas en estadística deben tener al elegir un estimador.

Búsqueda del mejor estimador

Un estadístico de la muestra dado no siempre es el mejor estimador de su parámetro poblacional correspondiente. Considere una población con distribución simétrica, en la que los valores de la mediana y de la media coinciden. En este caso, la media de la muestra sería un estimador *imparcial* de la mediana de la población. También, la media de la muestra sería un estimador *consistente* de la mediana de la población puesto que, al aumentar el tamaño de la muestra, el valor de la media de la muestra tenderá a acercarse bastante a la mediana de la población. Y la media de la muestra sería un estimador más *eficiente* de la mediana de la población que la mediana de la muestra misma, ya que en muestras grandes, la media de la muestra tiene un error estándar menor que la de la mediana de la muestra. Al mismo tiempo, la mediana de la muestra de una población con distribución simétrica sería un estimador imparcial y consistente de la media de la población, pero *no el más eficiente*, porque en muestras grandes su error estándar es mayor que el de la media de la muestra.

Ejercicios 7.1

- **7-1** ¿Cuales son las dos herramientas básicas que se utilizan al hacer inferencias estadísticas?
- **7-2** ¿Por qué los que toman decisiones a menudo miden muestras en lugar de medir poblaciones completas? ¿Cuál es la desventaja?
- **7-3** Explique una limitación que se presenta al hacer una estimación puntual, pero que no se presenta al hacer una estimación de intervalo.
- **7-4** ¿Qué es un estimador? ¿En qué se diferencia un estimador de una estimación?
- **7-5** Dé una lista de los criterios de un buen estimador y descríbalos brevemente.
- **7-6** ¿Qué papel juega la consistencia en la determinación del tamaño de la muestra?

7.2 Estimaciones puntuales

Uso de la media de la muestra para estimar la media de la población

La media de la muestra $\bar{x}$ es el mejor estimador de la media de la población μ. Es insesgada, consistente, el estimador más eficiente y, siempre y cuando la muestra sea suficientemente grande, su distribución muestral puede ser aproximada por medio de la distribución normal.

Si conocemos la distribución muestral de $\bar{x}$, podemos obtener conclusiones respecto a cualquier estimación que podamos hacer a partir de la información muestral. Considere el caso de una compañía de suministros clínicos que produce jeringas desechables. Cada jeringa está cubierta por una envoltura estéril que a su vez se empaca en grandes cajas de cartón corrugado. Debido al proceso de empaque, las cajas de cartón contienen distintas cantidades de jeringas. Como las jeringas se venden por pieza, la compañía necesita una estimación del número de piezas que hay por caja, para propósitos de facturación. Tomamos una muestra aleatoria de 35 cajas y registramos el número de jeringas contenidas en cada caja. La tabla 7-2 ilustra los resultados. Utilizando los conceptos del capítulo 3, podemos obtener la media de la muestra, $\bar{x}$, sumando todos los resultados, Σx, y dividiendo esta suma entre n, el número de cajas muestreadas:

Búsqueda de la media de la muestra

$$\bar{x} = \frac{\Sigma x}{n} \qquad [3\text{-}2]$$

Utilizando esta ecuación para resolver el problema, tenemos:

$$\bar{x} = \frac{3{,}570}{35}$$

$$= 102 \text{ jeringas}$$

Así, al usar la media de la muestra, $\bar{x}$ como estimador, la estimación puntual de la media de la población, μ, es 102 jeringas por caja. El precio de fabricación de cada jeringa hipodérmica desecha-

Tabla 7-2
Resultados obtenidos a partir de una muestra de 35 cajas (jeringas por caja)

101	103	112	102	98	97	93
105	100	97	107	93	94	97
97	100	110	106	110	103	99
93	98	106	100	112	105	100
114	97	110	102	98	112	99

ble es bastante bajo (alrededor de 25 centavos), de modo que tanto el comprador como el vendedor aceptarían esta estimación puntual como base para la facturación, y el fabricante puede ahorrarse el tiempo y el gasto de contar las jeringas contenidas en las cajas.

Estimación puntual de la varianza y la desviación estándar de la población

Uso de la desviación estándar de la muestra para estimar la desviación estándar de la población

Suponga que la administración de la compañía de suministros clínicos desea estimar la varianza y/o la desviación estándar de la distribución del número de jeringas empacadas por caja. El estimador más utilizado para estimar la desviación estándar de la población σ, es la desviación estándar de la muestra, s. Podemos calcular la desviación estándar de la muestra como lo hicimos en la tabla 7-3 y descubrir que es 6.01 jeringas.

Si en lugar de considerar

$$s^2 = \frac{\Sigma(x - \bar{x})^2}{n - 1}$$

¿Por qué el divisor es $n - 1$?

como nuestra varianza de la muestra, hubiéramos usado la ecuación:

$$s^2 = \frac{\Sigma(x - \bar{x})^2}{n}$$

el resultado habría tenido algo de *sesgo* como estimador de la varianza de la población; específicamente, hubiera tendido a ser demasiado bajo. Utilizar en el divisor $n - 1$, nos da un estimador imparcial de σ^2. En consecuencia, usaremos s^2 (según se define en la ecuación 3-17) y s (ecuación 3-18) para estimar σ^2 y σ.

Estimación puntual de la proporción de la población

Uso de la proporción de la muestra para estimar la proporción de la población

La proporción de unidades de una población dada que tiene una característica particular se denota por p. Si conocemos la proporción de unidades de una muestra que tiene la misma característica (denotada por $\hat{p}$, podemos utilizar esta $\hat{p}$ como estimador de p. Se puede demostrar que $\hat{p}$ tiene todas las características deseables analizadas; es insesgado (*no* sesgado), consistente, eficiente y suficiente.

Continuando con nuestro ejemplo del fabricante de suministros médicos, intentaremos hacer una estimación de la proporción de la población a partir de la proporción de la muestra. Suponga que la administración de la empresa desea estimar el número de cajas que llegarán dañadas a su destino por mal manejo en el traslado. Podemos verificar una muestra de 50 cajas a partir del punto de embarque hasta su arribo al punto de destino, y luego registrar la presencia o ausencia de daños. En este caso, si encontramos que la proporción de cajas dañadas en la muestra es 0.08, diríamos que:

$$\hat{p} = 0.08 \leftarrow \text{Proporción de la muestra dañada}$$

Y, debido a que la proporción de la muestra $\hat{p}$ es un estimador conveniente de la proporción de la población p, podemos estimar que la proporción de cajas dañadas de toda la población será también 0.08.

Tabla 7-3

Cálculo de la varianza y de la desviación estándar de la muestra para el número de jeringas por caja

Valores de x (jeringas por caja) (1)	x^2 (2)	Media de la muestra $\bar{x}$ (3)	$(x - \bar{x})$ (4) = (1) − (3)	$(x - \bar{x})^2$ (5) = (4)^2
101	10,201	102	−1	1
105	11,025	102	3	9
97	9,409	102	−5	25
93	8,649	102	−9	81
114	12,996	102	12	144
103	10,609	102	1	1
100	10,000	102	−2	4
100	10,000	102	−2	4
98	9,604	102	−4	16
97	9,409	102	−5	25
112	12,544	102	10	100
97	9,409	102	−5	25
110	12,100	102	8	64
106	11,236	102	4	16
110	12,100	102	8	64
102	10,404	102	0	0
107	11,449	102	5	25
106	11,236	102	4	16
100	10,000	102	−2	4
102	10,404	102	0	0
98	9,604	102	−4	16
93	8,649	102	−9	81
110	12,100	102	8	64
112	12,544	102	10	100
98	9,604	102	−4	16
97	9,409	102	−5	25
94	8,836	102	−8	64
103	10,609	102	1	1
105	11,025	102	3	9
112	12,544	102	10	100
93	8,649	102	−9	81
97	9,409	102	−5	25
99	9,801	102	−3	9
100	10,000	102	−2	4
99	9,801	102	−3	9
3,570	**365,368**	Suma de los cuadrados de todas las diferencias	$\Sigma(x - \bar{x})^2$ →	**1,228**

[3-17] $$s^2 = \frac{\Sigma x^2}{n-1} - \frac{n\bar{x}^2}{n-1}$$

$$= \frac{365{,}368}{34} - \frac{35(102)^2}{34}$$

$$= \frac{1{,}228}{34}$$

$$= 36.12$$

← o →

Suma de los cuadrados de las diferencias entre 34, el número de piezas de la muestra −1 (varianza de la muestra)

$$\frac{\Sigma(x - \bar{x})^2}{n-1} \rightarrow 36.12$$

[3-18] $$s = \sqrt{s^2} =$$

$$= \sqrt{36.12}$$

$$= 6.01 \text{ jeringas}$$

Desviación estándar de la muestra s

$$\sqrt{\frac{\Sigma(x - \bar{x})^2}{n-1}} \rightarrow 6.01 \text{ jeringas}$$

SUGERENCIAS Y SUPOSICIONES

Dejando de lado todas las definiciones, la razón para estudiar los estimadores es aprender acerca de las poblaciones mediante el muestreo, sin contar cada elemento de la población. Por supuesto, tampoco en este caso el viaje es gratis, y al decidir no contar todo, se pierde cierta exactitud. Los administradores desearían saber la exactitud que se logra cuando se hace un muestreo, y si usamos las ideas de este capítulo, podemos decírselo. Los estadísticos pueden establecer cómo se comporta el error estándar conforme aumenta o disminuye el tamaño de la muestra y los investigadores de mercados pueden determinar el costo de tomar más muestras o de hacerlas más grandes; pero deberá usar su propio juicio para combinar estos dos datos y tomar una decisiones *gerencial* correcta.

Ejercicios 7.2

Ejercicios de autoevaluación

EA 7-1 El Greensboro Coliseum estudia la posibilidad de ampliar su capacidad de asientos y necesita conocer tanto el número promedio de personas que asisten a los eventos como la variabilidad de este número. Los datos se refieren a la asistencia (en miles) a nueve eventos deportivos seleccionados al azar. Encuentre las estimaciones puntuales de la media y la varianza de la población de la que se tomó la muestra.

8.8 14.0 21.3 7.9 12.5 20.6 16.3 14.1 13.0

EA 7-2 La Autoridad para Distribución de Pizzas (ADP) ha desarrollado un buen negocio en Carrboro entregando órdenes de pizzas con prontitud. La ADP garantiza que sus pizzas se entregarán en 30 minutos o menos a partir del momento en que se toma el pedido y, si la entrega se retrasa, la pizza es gratis. El tiempo de entrega de cada pedido se registra en el "libro oficial de tiempo de pizza" (LOTP); el tiempo de entrega con retraso se registra como "30 minutos" en LOTP. Se enumeran 12 registros aleatorios del LOTP.

15.3 29.5 30.0 10.1 30.0 19.6
10.8 12.2 14.8 30.0 22.1 18.3

a) Encuentre la media de la muestra.
b) ¿De qué población se obtuvo esta muestra?
c) ¿Puede usarse esta muestra para estimar el tiempo promedio que toma a ADP entregar una pizza? Explique.

Aplicaciones

■ **7-7** A Joe Jackson, un meteorólogo que trabaja para la estación de televisión WDUL, le gustaría informar sobre la precipitación pluvial promedio para ese día en el noticiero de la tarde. Los datos siguientes corresponden a las mediciones de precipitación pluvial (en centímetros) para 16 años en la misma fecha, tomados al azar. Determine la precipitación pluvial media de la muestra.

0.47 0.27 0.13 0.54 0.00 0.08 0.75 0.06
0.00 1.05 0.34 0.26 0.17 0.42 0.50 0.86

■ **7-8** El National Bank of Lincoln quiere determinar el número de cajeros disponibles durante las horas pico del almuerzo los viernes. El banco ha recolectado datos del número de personas que entraron al banco los viernes de los últimos 3 meses entre las 11 A.M. y la 1 P.M. Utilice los siguientes datos para encontrar las estimaciones puntuales de la media y la desviación estándar de la población de donde se tomó la muestra.

242 275 289 306 342 385 279 245 269 305 294 328

■ **7-9** La empresa Electric Pizza está considerando la distribución a nivel nacional de su producto que ha tenido éxito a nivel local y para ello recabó datos de venta *pro forma*. Las ventas mensuales promedio (en miles de dólares) de sus 30 distribuidores actuales se listan a continuación. Tratando estos datos como a) una muestra y b) como una población, calcule la desviación estándar.

7.3	5.8	4.5	8.5	5.2	4.1
2.8	3.8	6.5	3.4	9.8	6.5
6.7	7.7	5.8	6.8	8.0	3.9
6.9	3.7	6.6	7.5	8.7	6.9
2.1	5.0	7.5	5.8	6.4	5.2

■ **7-10** En una muestra de 400 trabajadores textiles, 184 de ellos expresaron gran insatisfacción con el plan propuesto para modificar las condiciones de trabajo. Como el descontento de este grupo fue lo suficientemente fuerte para hacer que la administración de la fábrica considerara la reacción al plan como altamente negativa, tienen curiosidad de conocer la proporción del total de trabajadores en contra. Dé una estimación puntual de esta proporción.

■ **7-11** La red Amigos de los Videntes cobra $3 por minuto para conocer los secretos que pueden cambiar su vida. La red sólo cobra por minutos completos y redondea hacia arriba para beneficiar a la compañía. Así, una llamada de 2 minutos 10 segundos cuesta $9. Se da una lista de 15 cobros seleccionados al azar

3 9 15 21 42 30 6 9 6 15 21 24 32 9 12

a) Encuentre la media de la muestra.
b) Encuentre una estimación puntual de la varianza de la población.
c) ¿Puede esta muestra usarse para estimar la duración promedio de una llamada? Si es así, ¿cuál es la estimación? Si no, ¿qué se puede estimar con esta muestra?

Soluciones a los ejercicios de autoevaluación

EA 7-1 $\Sigma x^2 = 2003.65 \qquad \Sigma x = 128.5 \qquad n = 9$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{128.5}{9} = 14.2778 \text{ miles de personas}$$

$$s^2 = \frac{1}{n-1}(\Sigma x^2 - n\bar{x}^2) = \frac{2003.65 - 9(14.2778)^2}{8}$$

$$= 21.119 \text{ (miles de personas)}^2$$

EA 7-2 a) $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{247.7}{12} = 20.225$ minutos.

b) La población de tiempos registrados en el LOTP.
c) No, no se puede. Debido a que el tiempo de entrega mayor que 30 minutos se registra como 30 minutos, usar estos datos subestimará en forma consistente el promedio del tiempo de entrega.

7.3 Estimaciones de intervalo: conceptos básicos

El propósito de tomar muestras es conocer más acerca de una población. Podemos calcular esta información a partir de las muestras como estimaciones *puntuales*, que acabamos de analizar, o como estimaciones de *intervalo*, que son el tema del resto de este capítulo. **Una estimación de intervalo describe un rango de valores dentro del cual es posible que esté un parámetro de la población.**

Iniciamos con, la estimación puntual

Suponga que el director de estudios de mercado de una fábrica de refacciones automotrices necesita hacer una estimación de la vida promedio de las baterías para automóvil que produce su compañía. Seleccionamos una muestra aleatoria de 200 baterías, registramos el nombre y dirección de los propietarios de los automóviles, como están en los registros de ventas, y entrevistamos a estas personas con respecto a la duración de la batería de su automóvil. Nuestra muestra de 200 usuarios tiene una vida media de las baterías de 36 meses. Si utilizamos la estimación puntual de la media de

la muestra $\bar{x}$ como el mejor estimador de la media de la población μ, informaríamos que la vida media de las baterías de la empresa es 36 meses.

Búsqueda del error probable de esta estimación

Pero el director también pide una conclusión acerca de la incertidumbre que acompañará a esta estimación; es decir, una afirmación acerca del intervalo dentro del cual es probable que esté la media de la población desconocida. Para proporcionar tal afirmación, necesitamos encontrar el *error estándar de la media.*

En el capítulo 6 aprendimos que si seleccionamos y graficamos un número grande de medias de muestras de una población, la distribución de estas medias se aproximará a la curva normal. Además, la media de las medias muestrales será la misma que la media de la población. Nuestro tamaño de muestra de 200 baterías es suficientemente grande para poder aplicar el teorema central del límite; como se hizo de manera gráfica en la figura 7-1. Para medir la extensión, o dispersión, de nuestra distribución de medias muestrales, podemos utilizar la siguiente fórmula* y calcular el error estándar de la media:

Error estándar de la media para una población infinita → $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$ ← Desviación estándar de la población [6-1]

Suponga que ya se estimó la desviación estándar de la población de baterías y se informó que es 10 meses. Con esta desviación estándar y la primera ecuación del capítulo 6, podemos calcular el error estándar de la media:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \qquad [6\text{-}1]$$

$$= \frac{10}{\sqrt{200}}$$

$$= \frac{10}{14.14}$$

$$= 0.707 \text{ meses} \leftarrow \text{Un error estándar de la media}$$

Obtención de la estimación de intervalo

Ahora, podemos informar al director que nuestra estimación de la vida útil de las baterías de la compañía es 36 meses y que el error estándar que acompaña a esta estimación es 0.707. En otras palabras, la vida útil real para todas las baterías *puede* estar en alguna parte de la estimación de intervalo comprendida entre 35.293 y 36.707 meses. Esto es útil pero no es suficiente información para el director. Necesitamos calcular la posibilidad de que la duración real de las baterías esté en este intervalo *o* en otros intervalos de diferentes anchos que podamos escoger, $\pm 2\sigma$ (2×0.707), $\pm 3\sigma$ (3×0.707), y así sucesivamente.

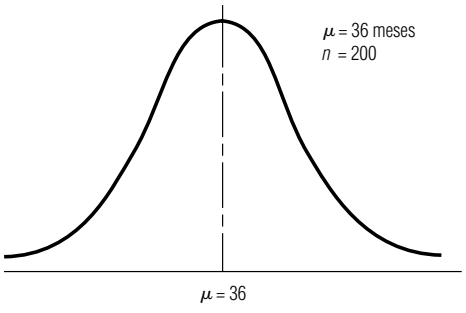


FIGURA 7-1

Distribución muestral de la media para muestras de 200 baterías

* No se usó el multiplicador de población finita para calcular el error estándar de la media porque la población de baterías es lo suficientemente grande para considerarla infinita.

Probabilidad de que el verdadero parámetro poblacional caiga dentro de la estimación del intervalo

Para empezar a resolver este problema, debemos repasar las partes importantes del capítulo 5. Trabajamos con la distribución normal de probabilidad y aprendimos que porciones específicas del área bajo la curva normal están localizadas entre más-menos cierto número de desviaciones estándar a partir de la media. En la figura 5-12 vimos cómo relacionar estas porciones con probabilidades específicas.

Búsqueda de la probabilidad de que la media caiga en esta estimación del intervalo

Afortunadamente, podemos aplicar estas propiedades al error estándar de la media y afirmar lo siguiente acerca del **rango** de valores que se utilizaron para hacer una estimación de intervalo en nuestro problema de las baterías.

La probabilidad es 0.955 de que la media de una muestra de 200 baterías esté dentro de ± 2 errores estándar de la media de la población. Dicho de manera diferente, el 95.5% de todas las medias muestrales está dentro de ± 2 errores estándar de μ y, en consecuencia, μ **está dentro de ± 2 errores estándar del 95.5% de todas las medias muestrales**. Teóricamente, si seleccionamos 1,000 muestras al azar de una población dada y luego construimos un intervalo de ± 2 errores estándar alrededor de la media de cada una de esas muestras, cerca de 955 de estos intervalos incluirán a la media de la población. De manera parecida, la probabilidad de que la media de la muestra esté dentro de ± 1 error estándar de la media de la población es 0.683, y así sucesivamente. Este concepto teórico es fundamental para nuestro estudio sobre la construcción de intervalos y la inferencia estadística. La figura 7-2 ilustra el concepto de manera gráfica e indica cinco de esos intervalos. Únicamente el intervalo construido alrededor de la media de la muestra $\bar{x}_4$ no contiene a la media de la población. En palabras, los estadísticos describirían las estimaciones de intervalos representadas en la figura 7-2 como sigue: "La media de la población μ, estará localizada dentro de ± 2 errores estándar de la media muestral el 95.5% de las veces."

En lo que concierne a cualquier intervalo particular de la figura 7-2, éste contiene a la media de la población o no la contiene, pues la media de la población es un parámetro fijo. Como sabemos que el 95.5% de todas las muestras el intervalo contendrá a la media de la población, decimos que hay 95.5% de confianza de que el intervalo contenga a la media de la población.

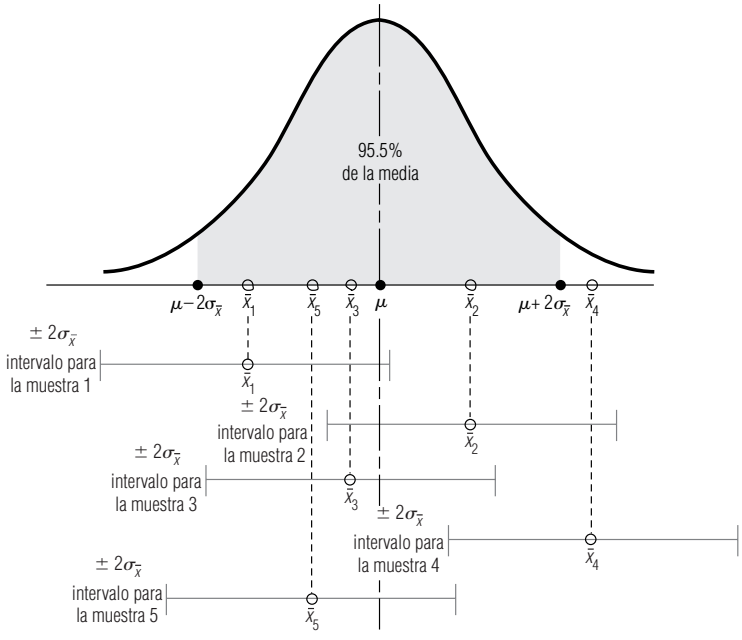


FIGURA 7-2

Cierto número de intervalos construidos alrededor de las medias muestrales; todos, excepto uno, incluyen a la media de la población

Una estimación más útil de la vida de las baterías

Con la aplicación de lo anterior al ejemplo de las baterías, podemos dar un informe al director. Nuestra mejor estimación de la vida útil de las baterías de la compañía es 36 meses, *y* tenemos 68.3% de confianza de que la vida útil se encuentra en el intervalo que abarca de 35.293 a 36.707 meses ($36 \pm 1\sigma_{\bar{x}}$). Similarmente, tenemos 95.5% de confianza de que la duración caiga dentro del intervalo comprendido entre 34.586 y 37.414 meses ($36 \pm 2\sigma_{\bar{x}}$), y tenemos el 99.7% de confianza de que la vida útil de una batería estará dentro del intervalo que va de 33.879 a 38.121 meses ($36 \pm 3\sigma_{\bar{x}}$).

SUGERENCIAS Y SUPOSICIONES

Cada vez que se hace una estimación existe un error implícito en ella. Para que las personas lo entiendan, es una práctica común describirlo con una afirmación como "nuestra mejor estimación de la vida de estas llantas es 40,000 millas y tenemos una seguridad del 90% de que la vida estará entre 35,000 y 45,000 millas". Pero si su jefe quiere saber cuál es la vida promedio exacta de un conjunto de llantas, y no supiera de muestreo, tendría que observar cientos de miles de conjuntos de llantas hasta que se desgastaran, y después calcular cuánto duraron en promedio. Advertencia: incluso en este caso estaría haciendo un muestreo porque es imposible observar y medir todos los juegos de llantas que están en uso. Es mucho menos costoso y más rápido usar el muestreo para encontrar la respuesta. Si entiende las estimaciones, puede decirle a su jefe qué riesgos implica usar una muestra para estimar la vida útil real de la llanta.

Ejercicios 7.3

Ejercicios de autoevaluación

EA 7-3 Para una población con una varianza conocida de 185, una muestra de 64 individuos lleva a 217 como estimación de la media.

a) Encuentre el error estándar de la media.
b) Establezca una estimación de intervalo que incluya la media de la población el 68.3% del tiempo.

EA 7-4 Eunice Gunterwal es una ahorradora estudiante de licenciatura de la universidad del estado que está interesada en comprar un auto usado. Selecciona al azar 125 anuncios y ve que el precio promedio de un auto en esta muestra es $3,250. Eunice sabe que la desviación estándar de los precios de los autos usados en esta ciudad es $615.

a) Establezca una estimación de intervalo para el precio promedio de un automóvil de manera que Eunice tenga una seguridad del 68.3% de que la media de la población está dentro de este intervalo.
b) Establezca una estimación de intervalo para el precio promedio de un auto de modo que la señorita Gunterwal tenga el 95.5% de certeza de que la media de la población está dentro de este intervalo.

Conceptos básicos

■ **7-12** De una población que se sabe tiene una desviación estándar de 1.4, se toma una muestra de 60 individuos. Se encuentra que la media de esta muestra es 6.2.

a) Encuentre el error estándar de la media.
b) Construya una estimación de intervalo alrededor de la media de la muestra, utilizando un error estándar de la media.

■ **7-13** De una población con desviación estándar conocida de 1.65, una muestra de 32 elementos dio como resultado 34.8 como estimación de la media.

a) Encuentre el error estándar de la media.
b) Calcule un intervalo estimado que incluya la media de la población el 99.7% del tiempo.

Aplicaciones

■ **7-14** La Universidad de Carolina del Norte está llevando a cabo un estudio sobre el peso promedio de los adoquines que conforman los andadores del campus. Se envía a algunos trabajadores a desenterrar y pesar

una muestra de 421 adoquines, y el peso promedio de la muestra resulta ser 14.2 libras. Todo mundo sabe que la desviación estándar del peso de un adoquín es 0.8 libras.

a) Encuentre el error estándar de la media.
b) ¿Cuál es el intervalo alrededor de la media de la muestra que incluirá la población de la media el 95.5% de las veces?

■ **7-15** Debido a que el dueño del restaurante recientemente abierto, El Refugio del Bardo ha tenido dificultades al estimar la cantidad de comida que debe preparar cada tarde, ha decidido determinar el número medio de clientes a los que atiende cada noche. Seleccionó una muestra de 30 noches que le arrojaron una media de 71 clientes. Se llegó a la conclusión de que la desviación estándar de la población es 3.76.

a) Dé una estimación de intervalo que tenga el 68.3% de probabilidad de incluir a la media de la población.
b) Dé una estimación de intervalo que tenga el 99.7% de probabilidad de incluir a la media de la población.

■ **7-16** La administradora del puente Neuse River está preocupada acerca de la cantidad de automóviles que pasan sin pagar por las casetas de cobro automáticas del puente, y está considerando cambiar la manera de cobrar, si el cambio permite solucionar el problema. Muestreó al azar 75 horas para determinar la tasa de violación. El número promedio de violaciones por hora fue 7. Si se sabe que la desviación estándar de la población es 0.9, estime un intervalo que tenga el 95.5% de probabilidad de contener a la media verdadera.

■ **7-17** Gwen Taylor, administradora de los departamentos WilowWood, desea informar a los residentes potenciales cuánta energía eléctrica pueden esperar usar durante el mes de agosto. Selecciona 61 residentes aleatorios y descubre que su consumo promedio en agosto es 894 kilowatts hora (kwh). Gwen piensa que la varianza del consumo es alrededor de 131 $(\text{kwh})^2$.

a) Establezca una estimación de intervalo para el consumo promedio de energía eléctrica en el mes de agosto para que Gwen pueda tener una seguridad del 68.3% de que la media verdadera de la población está dentro de este intervalo.
b) Repita la parte a) para una certeza del 99.7%.
c) Si el precio por kilowatt es \$0.12, ¿dentro de qué intervalo puede Gwen estar 68.3% segura que caerá el costo promedio de agosto por consumo de electricidad?

■ **7-18** La Junta Directiva de Escuelas Estatales del condado Pesimismo considera que su tarea más importante es mantener el tamaño promedio de los grupos de sus escuelas menor que el tamaño promedio de los grupos de Optimismo, el condado vecino. Dee Marks, la superintendente de escuelas de Pesimismo, acaba de recibir información confiable que indica que el tamaño del grupo promedio en Optimismo este año es 30.3 estudiantes. Todavía no tiene los datos correspondientes de los 621 grupos de su propio sistema escolar, de modo que Dee se ve forzada a basar sus cálculos en los 76 grupos que han informado acerca de su tamaño de grupo, que producen un promedio de 29.8 estudiantes. Dee sabe que el tamaño de grupo de las escuelas de Pesimismo tiene una distribución con media desconocida y una desviación estándar de 8.3 estudiantes. Suponiendo que la muestra de 76 estudiantes que tiene la señorita Marks es una muestra aleatoria de la población de los grupos del condado Pesimismo:

a) Encuentre un intervalo en el cual Dee Marks pueda tener el 95.5% de certeza de que contendrá a la media real.
b) ¿Usted cree que la señora Dee ha conseguido su objetivo?

Soluciones a lo ejercicios de autoevaluación

EA 7-3 $\sigma^2 = 185 \quad \sigma = \sqrt{185} = 13.60 \quad n = 64 \quad \bar{x} = 217$

a) $\sigma_{\bar{x}} = \sigma/\sqrt{n} = 13.60/\sqrt{64} = 1.70$
b) $\bar{x} \pm \sigma_{\bar{x}} = 217 \pm 1.70 = (215.3, 218.7)$

EA 7-4 $\sigma = 615 \quad n = 125 \quad \bar{x} = 3{,}250 \quad \sigma_{\bar{x}} = \sigma/\sqrt{n} = 615/\sqrt{125} = 55.01$

a) $\bar{x} \pm \sigma_{\bar{x}} = 3{,}250 \pm 55.01 = (\$3{,}194.99, \$3{,}305.01)$
b) $\bar{x} \pm 2\sigma_{\bar{x}} = 3{,}250 \pm 2(55.01) = 3{,}250 \pm 110.02 = (\$3{,}139.98, \$3{,}360.02)$

7.4 Estimaciones de intervalo e intervalos de confianza

Al utilizar estimaciones de intervalo no nos estamos limitando a ± 1, 2 y 3 errores estándar. De acuerdo con la tabla 1 del apéndice, ± 1.64 errores estándar, por ejemplo, incluyen aproximadamente el

90% del área bajo la curva y, así, 0.4495 del área a ambos lados de la media en una distribución normal. De manera parecida, ± 2.58 errores estándar incluyen alrededor de 99% del área o el 49.51% a cada lado de la media.

Definición de *nivel de confianza*

En estadística, la probabilidad que asociamos con una estimación de intervalo se conoce como *nivel de confianza.* Esta probabilidad indica qué tanta confianza tenemos de que la estimación de intervalo incluya al parámetro de población. Una probabilidad más alta implica una mayor confianza. En la estimación, los niveles de confianza que se utilizan con más frecuencia son 90, 95 y 99%, pero somos libres de aplicar *cualquier* nivel de confianza. En la figura 7-2, por ejemplo, utilizamos un nivel de confianza del 95.5%.

El *intervalo de confianza* es el rango de la estimación que estamos haciendo. Si informamos que tenemos el 90% de confianza de que la media de la población de ingresos de las personas que viven en una cierta comunidad está entre \$8,000 y \$24,000, entonces el rango \$8,000-\$24,000 es nuestro intervalo de confianza. A menudo, sin embargo, expresaremos el intervalo de confianza en términos de errores estándar, más que con valores numéricos. Así, expresaremos los intervalos de confianza de esta forma: $\bar{x} \pm 1.64\sigma_{\bar{x}}$, donde:

$$\bar{x} + 1.64\sigma_{\bar{x}} = \text{límite superior del intervalo de confianza}$$

$$\bar{x} - 1.64\sigma_{\bar{x}} = \text{límite inferior del intervalo de confianza}$$

Entonces, los *límites de confianza* son los límites superior e inferior del intervalo de confianza. En este caso, $\bar{x} + 1.64\sigma_{\bar{x}}$ se conoce como *límite superior de confianza (LSC)* y $\bar{x} - 1.64\sigma_{\bar{x}}$ es el *límite inferior de confianza (LIC).*

Relación entre nivel de confianza e intervalo de confianza

Podría pensarse que deberíamos utilizar un alto nivel de confianza, como 99%, en todos los problemas sobre estimaciones. Después de todo, parece ser que un alto nivel de confianza significa un alto grado de precisión en lo que a la estimación concierne. En la práctica, sin embargo, altos niveles de confianza producen intervalos de confianza grandes, y éstos, de hecho, dan estimaciones bastante imprecisas.

Considere, por ejemplo, el caso de un cliente de una tienda de electrodomésticos que pregunta sobre la entrega de una nueva lavadora de ropa. En la tabla 7-4 presentamos varias preguntas que el cliente podría hacer y las respuestas probables. Esta tabla indica la relación directa que existe entre el nivel de confianza y el intervalo de confianza de cualquier estimación. A medida que el cliente va estableciendo un intervalo de confianza cada vez más estrecho, el administrador de la tienda consiente en un nivel de confianza cada vez más bajo. Note, también, que cuando el intervalo de confianza es demasiado amplio, como en el caso en que la entrega tarda un año, la estimación puede tomar un valor real muy pequeño, a pesar, incluso, de que el administrador le da un nivel de confianza del 99% a dicha estimación. De manera parecida, si el intervalo de confianza es muy reducido (¿Llegará la nueva lavadora a mi casa antes que yo?), la estimación está asociada a un nivel de confianza tan bajo (l%) que cuestionamos su valor.

Uso de muestreo y estimación de intervalos de confianza

Estimación a partir de una sola muestra

En nuestro análisis de los conceptos básicos de la estimación de intervalos, particularmente en la figura 7-2, describimos muestras de una población tomadas de manera repetida con el fin de estimar un parámetro. Mencionamos, también, la selección de un gran número de medias muestrales de una población. En la práctica, sin embargo, a menudo resulta difícil o costoso tomar más de una muestra de una población. Con base en una sola muestra estimamos el parámetro de la población. Debemos tener cuidado, entonces, en la interpretación de los resultados de este proceso.

Suponga que calculamos, en el ejemplo de las baterías y a partir de una sola muestra, el intervalo de confianza y el nivel de confianza: "Tenemos el 95% de confianza de que la vida media de las

Tabla 7-4
Ilustración de la relación entre nivel de confianza e intervalo de confianza

Pregunta del cliente	Respuesta del administrador de la tienda	Nivel de confianza implicado	Intervalo de confianza implicado
¿Llegará la lavadora antes de un año?	Tengo la absoluta certeza de ello.	Mayor que 99%	Un año
¿Me entregarán la lavadora antes de un mes?	Estoy casi seguro que la recibirá en este mes.	Al menos 95%	Un mes
¿Me entregarán la lavadora antes de una semana?	Estoy bastante seguro de que saldrá en esta semana.	Alrededor del 80%	Una semana
¿Tendré la lavadora en mi casa mañana?	No tengo la certeza de poder hacerlo.	Alrededor del 40%	Un día
¿Llegará la nueva lavadora a mi casa antes que yo?	Hay una pequeña posibilidad.	Cercano al 1%	Una hora

baterías de la población se encuentra entre 30 y 42 meses." **Esta afirmación no significa que se tiene 0.95 de probabilidad de que la vida media de todas las baterías caiga dentro del intervalo establecido para esta muestra. Más bien, indica que si seleccionamos muchas muestras aleatorias del mismo tamaño y calculamos un intervalo de confianza para cada una de esas muestras, entonces en alrededor del 95% de los casos la media de la población caerá dentro de dicho intervalo.**

SUGERENCIAS Y SUPOSICIONES

Nada es gratis en lo que respecta a niveles e intervalos de confianza. Cuando obtiene más de uno, deberá tener menos del otro. Es recomendable, para comprender esta importante relación, que regrese a la tabla 7-4. Si desea que la estimación del tiempo de entrega tenga una exactitud *perfecta* del (100%), deberá sacrificar precisión en el intervalo de confianza y aceptar una promesa amplia de tiempo de entrega ("en algún momento del año"). Por otro lado, si no le preocupa la exactitud de la estimación, su personal de entrega podría decir "tengo una seguridad del 1% de que podemos entregarle en menos de 1 hora". *No se puede* tener las dos cosas al mismo tiempo.

Ejercicios 7.4

Ejercicios de autoevaluación

EA 7-5 Dados los siguientes niveles de confianza, exprese los límites inferior y superior del intervalo de confianza para estos niveles en términos de $\bar{x}$ y $\sigma_{\bar{x}}$.

a) 54%.
b) 75%.
c) 94%.
d) 98%.

Conceptos básicos

7-19 Defina el nivel de confianza para una estimación de intervalo.

7-20 Defina el intervalo de confianza.

7-21 Suponga que desea utilizar un nivel de confianza del 80%. Dé el límite superior del intervalo de confianza en términos de la media de la muestra, $\bar{x}$, y del error estándar, $\sigma_{\bar{x}}$.

7-22 ¿De qué forma podría una estimación ser menos significativa debido a

a) un alto nivel de confianza?
b) un estrecho nivel de confianza?

7-23 Suponga que se toma una muestra de 50 elementos de una población con desviación estándar de 27, y que la media de la muestra es 86.

a) Establezca una estimación de intervalo para la media de la población que tenga el 95.5% de certeza de incluir a la media verdadera de la población.
b) Suponga, ahora, que el tamaño de la muestra es 5,000 elementos. Establezca un intervalo para la media de la población que tenga el 95.5% de certeza de incluir a la media verdadera de la población.
c) ¿Por qué la estimación del inciso a) sería preferible a la del inciso b)? ¿Por qué la estimación del inciso b) sería mejor que la del inciso a)?

7-24 El nivel de confianza para una estimación, ¿está basado en el intervalo obtenido a partir de una sola muestra?

7-25 Dados los siguientes niveles de confianza, exprese los límites inferior y superior del intervalo de confianza en términos de $\bar{x}$ y de $\sigma_{\bar{x}}$.

a) 60%.
b) 70%.
c) 92%.
d) 96%.

Aplicaciones

7-26 Steve Klippers, dueño de la peluquería Steve´s, se ha formado una buena reputación entre los residentes de Cullowhee. Cuando un cliente entra a su establecimiento, Steve grita los minutos que el cliente deberá esperar antes de que se le atienda. El único estadístico del pueblo, después de frustrarse por las poco precisas estimaciones puntuales de Steve, ha determinado que el tiempo de espera real de cualquier cliente está distribuido normalmente con una media igual a la estimación de Steve en minutos y una desviación estándar igual a 5 minutos divididos entre la posición del cliente en la fila de espera. Ayude a los clientes de Steve´s a establecer intervalos con el 95% de probabilidad para las situaciones siguientes:

a) El cliente es el segundo en la fila y la estimación de Steve es 25 minutos.
b) El cliente es el tercero y la estimación de Steve es 15 minutos.
c) El cliente es el quinto de la fila, y la estimación de Steve es 38 minutos.
d) El cliente es el primero de la fila, y la estimación de Steve es 20 minutos. ¿Qué diferencia existe entre estos intervalos y los intervalos de confianza?

Soluciones a los ejercicios de autoevaluación

EA 7-5 a) $\bar{x} \pm 0.74\sigma_{\bar{x}}$. b) $\bar{x} \pm 1.15\sigma_{\bar{x}}$. c) $\bar{x} \pm 1.88\sigma_{\bar{x}}$. d) $\bar{x} \pm 2.33\sigma_{\bar{x}}$.

7.5 Cálculo de estimaciones de intervalo de la media a partir de muestras grandes

Búsqueda de un intervalo de confianza del 95%

Un mayorista de refacciones automotrices necesita una estimación de la vida media que puede esperar de los limpiadores de parabrisas en condiciones normales de manejo. La administración de la empresa ya ha determinado que la desviación estándar de la vida útil de la población es 6 meses. Suponga que seleccionamos una sola muestra aleatoria de 100 limpiadores, tomamos los datos referentes a su vida útil y obtenemos los siguientes resultados:

$n = 100$ ← Tamaño de la muestra
$\bar{x} = 21$ meses ← Media de la muestra
$\sigma = 6$ meses ← Desviación estándar de la población

Desviación estándar de la población

Como el distribuidor utiliza decenas de miles de limpiadores al año, nos pide que encontremos una estimación de intervalo con un nivel de confianza del 95%. El tamaño de la muestra es mayor que 30, de modo que el teorema central del límite nos permite usar la distribución normal como distri-

bución de muestreo, aun cuando nuestra población no tenga distribución normal. Calculamos el error estándar de la media con la ecuación 6-1:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \qquad [6\text{-}1]$$

$$= \frac{6 \text{ meses}}{\sqrt{100}}$$

$$= \frac{6}{10}$$

$$= 0.6 \text{ meses} \leftarrow \text{Error estándar de la media para una población infinita}$$

Cálculo de los límites de confianza

A continuación consideraremos el nivel de confianza con el cual estamos trabajando. Como un nivel del 95% de confianza incluirá el 47.5% del área que se encuentra a ambos lados de la media de la distribución de muestreo, podemos buscar en el cuerpo de la tabla 1 del apéndice el valor correspondiente a 0.475. Descubrimos que 0.475 del área bajo la curva normal está contenida entre la media y un punto situado a 1.96 errores estándar a la derecha de la media. Por consiguiente, sabemos que (2)(0.475) = 0.95 del área está localizada entre ±1.96 errores estándar de la media y que nuestros límites de confianza son:

$$\bar{x} + 1.96\sigma_{\bar{x}} \leftarrow \text{Límite superior de confianza}$$

$$\bar{x} - 1.96\sigma_{\bar{x}} \leftarrow \text{Límite inferior de confianza}$$

Luego sustituimos valores numéricos en estas dos expresiones:

$$\bar{x} + 1.96\sigma_{\bar{x}} = 21 \text{ meses} + 1.96(0.6 \text{ meses})$$
$$= 21 + 1.18 \text{ meses}$$
$$= 22.18 \text{ meses} \leftarrow \text{Límite superior de confianza}$$
$$\bar{x} - 1.96\,\sigma_{\bar{x}} = 21 \text{ meses} - 1.96(0.6 \text{ meses})$$
$$= 21 - 1.18 \text{ meses}$$
$$= 19.82 \text{ meses} \leftarrow \text{Límite inferior de confianza}$$

Nuestra conclusión

Ahora podemos informar que estimamos la vida media de la población de limpiadores de parabrisas entre 19.82 y 22.18 meses con un 95% de confianza.

Cuando no se conoce la desviación estándar de la población

Búsqueda de un intervalo de confianza del 90%

Un problema más complejo de estimación de intervalo proviene del departamento de servicio social de una dependencia gubernamental local. El departamento está interesado en estimar el ingreso medio anual de 700 familias que viven en una sección de cuatro manzanas de una comunidad. Tomamos una muestra aleatoria simple y encontramos los siguientes resultados:

$$n = 50 \leftarrow \text{Tamaño de muestra}$$

$$\bar{x} = \$11{,}800 \leftarrow \text{Media de la muestra}$$

$$s = \$950 \leftarrow \text{Desviación estándar de la muestra}$$

El departamento nos pide que calculemos una estimación de intervalo del ingreso anual medio de las 700 familias, de modo que pueda tener el 90% de confianza de que la media de la población

se encuentra dentro de ese intervalo. El tamaño de la muestra es mayor que 30, de manera que, de nuevo, el teorema central del límite nos permite utilizar la distribución normal como la distribución de muestreo.

Estimación de la desviación estándar de la población

Observe que una parte de este problema es diferente de los ejemplos anteriores; *no* conocemos la desviación estándar de la población y, por tanto, utilizaremos la desviación estándar de la muestra para estimar la *desviación estándar de la población*:

Estimación de la desviación estándar de la población

Estimación de la desviación estándar de la población $\longrightarrow \hat{\sigma} = s = \sqrt{\dfrac{\Sigma(x - \bar{x})^2}{n - 1}}$ [7-1]

El valor de $950.00 es nuestra estimación de la desviación estándar de la población. El símbolo para representar este valor *estimado* es $\hat{\sigma}$, que se conoce como *sigma gorro.*

Ahora podemos estimar el error estándar de la media. Como tenemos un tamaño de población finito y nuestra muestra constituye más del 5% de la población, utilizaremos la fórmula para derivar el error estándar de la media de poblaciones finitas:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \times \sqrt{\frac{N - n}{N - 1}} \qquad [6\text{-}3]$$

Estimación del error estándar de la media

Ya que estamos calculando el error estándar de la media mediante una *estimación* de la desviación estándar de la población, volvemos a escribir esta ecuación de modo que los símbolos sean correctos:

Estimación del error estándar de la media de una población finita

Símbolo que indica un valor estimado; Estimación de la desviación estándar de la población

$$\hat{\sigma}_{\bar{x}} = \frac{\hat{\sigma}}{\sqrt{n}} \times \sqrt{\frac{N - n}{N - 1}} \qquad [7\text{-}2]$$

Continuando con nuestro ejemplo, encontramos que $\hat{\sigma}_{\bar{x}} = \dfrac{\$950.00}{\sqrt{50}} \times \sqrt{\dfrac{700 - 50}{700 - 1}}$

$$= \frac{\$950.00}{7.07}\sqrt{\frac{650}{699}}$$

$$= (\$134.37)(0.9643)$$

$$= \$129.57$$ ← Estimación del error estándar de la media de una población finita (derivada de una *estimación* de la desviación estándar de la población)

En seguida consideramos el nivel de confianza del 90%, que incluiría el 45% del área que se encuentra a ambos lados de la media de la distribución de muestreo. Si observamos la tabla 1 del apéndice y buscamos el valor correspondiente a 0.45, encontramos que aproximadamente 0.45 del área bajo la curva normal está localizada entre la media y un punto alejado de ésta 1.64 errores estándar. En consecuencia, el 90% del área está localizada entre ± 1.64 errores estándar de la media, y nuestros límites de confianza son:

$$\bar{x} + 1.64\hat{\sigma}_{\bar{x}} = \$11{,}800 + 1.64\ (\$129.57)$$
$$= \$11{,}800 + \$212.50$$
$$= \$12{,}012.50 \leftarrow \text{Límite de confianza superior}$$
$$\bar{x} - 1.64\hat{\sigma}_{\bar{x}} = \$11{,}800 - 1.64(\$129.57)$$
$$= \$11{,}800 - \$212.50$$
$$= \$11{,}587.50 \leftarrow \text{Límite de confianza inferior}$$

Nuestra conclusión

El informe que podríamos dar al departamento de servicio social sería: "Con una confianza del 90%, estimamos que el ingreso anual promedio de las 700 familias que viven en una sección de cuatro manzanas se encuentra entre $11,587.50 y $12,012.50."

SUGERENCIAS Y SUPOSICIONES

Es sencillo entender cómo comenzar a resolver estos ejercicios si regresa a la figura 7-2 un momento. Cuando alguien establece un nivel de confianza, se refiere al área sombreada de la figura, que se define por cuántas $\sigma_{\bar{x}}$ (errores estándar o desviaciones estándar de la distribución de medias muestrales) hay a cada lado de la media. La tabla 1 del apéndice convierte cualquier nivel de confianza deseado en errores estándar. Como se cuenta con la información necesaria para calcular *un* error estándar, es posible calcular los puntos terminales del área sombreada; éstos son los límites del intervalo de confianza. Recuerde que cuando no se conoce la dispersión de la población (la desviación estándar de la población) puede usar la ecuación 7-1 para estimarla.

Ejercicios 7.5

Ejercicios de autoevaluación

EA 7-6 Se toma una muestra de 60 individuos a partir de una población de 540. De esta muestra, se encuentra que la media es 6.2 y la desviación estándar es 1.368.

a) Encuentre la estimación del error estándar de la media.
b) Construya un intervalo del 96% de confianza para la media.

EA 7-7 En una prueba de seguridad automovilística realizada por el Centro de Investigación Carretera de Carolina del Norte, la presión promedio de las llantas para una muestra de 62 llantas fue 24 libras por pulgada cuadrada y la desviación estándar fue 2.1 libras por pulgada cuadrada.

a) ¿Cuál es la desviación estándar estimada para esta población? (Existen cerca de un millón de automóviles registrados en Carolina del Norte).
b) Calcule el error estándar estimado de la media.

Conceptos básicos

c) Construya un intervalo de confianza del 95% para la media de la población.

■ **7-27** El gerente de la división de bombillas de la Cardinal Electric debe estimar el número promedio de horas que durarán los focos fabricados por cada una de las máquinas. Fue elegida una muestra de 40 focos de la máquina A y el tiempo promedio de funcionamiento fue 1,416 horas. Se sabe que la desviación estándar de la duración es 30 horas.

a) Calcule el error estándar de la media.
b) Construya un intervalo de confianza del 90% para la media de la población.

■ **7-28** Después de recolectar una muestra de 250 elementos de una población con una desviación estándar conocida de 13.7, se encuentra que la media es 112.4.

a) Encuentre un intervalo de confianza del 95% para la media.
b) Encuentre un intervalo de confianza del 99% para la media.

Aplicaciones

7-29 La enfermera de la secundaria de Westview está interesada en conocer la estatura promedio de los estudiantes del último año, pero no tiene suficiente tiempo para examinar los registros de los 430 estudiantes. Por ello, selecciona 48 al azar y encuentra que la media de la muestra es 64.5 pulgadas y la desviación estándar es 2.3 pulgadas.

a) Encuentre la estimación del error estándar de la media.
b) Construya un intervalo de confianza del 90% para la media.

7-30 Jon Jackobsen, un pasante de posgrado muy dedicado, acaba de terminar una primera versión de su tesis de 700 páginas. Jon mecanografió el trabajo por sí mismo y está interesado en conocer el número promedio de errores tipográficos por página, pero no quiere leer todo el documento. Como sabe algo acerca de estadística para la administración, Jon leyó 40 páginas seleccionadas de manera aleatoria y encontró que el promedio de errores tipográficos por página fue 4.3 y la desviación estándar de la muestra fue 1.2 errores por página.

a) Calcule el error estándar estimado de la media.
b) Calcule un intervalo de confianza del 90% para el número promedio verdadero de errores por página en su trabajo.

7-31 La Autoridad para la Televisión por Cable de Nebraska (ATCN) realizó una prueba para determinar el tiempo que las personas pasan frente al televisor por semana. La ATCN encuestó a 84 suscriptores y encontró que el número promedio de horas que ven televisión por semana es 11.6 horas con una desviación estándar de 1.8 horas.

a) ¿Cuál es la desviación estándar de la población estimada para esta población? (Existen cerca de 95,000 personas con televisión por cable en Nebraska.)
b) Calcule el error estándar estimado de la media.
c) Construya un intervalo de confianza del 98% para la media de la población.

7-32 Joel Friedlander es un corredor de la Bolsa de Valores de Nueva York y tiene curiosidad acerca del tiempo que transcurre entre la colocación de una orden de venta y su ejecución. Joel hizo un muestreo de 45 órdenes y encontró que el tiempo medio para la ejecución fue 24.3 minutos, con una desviación estándar de 3.2 minutos. Ayude a Joel con la construcción de un intervalo de confianza del 95% para el tiempo medio para la ejecución de una orden.

7-33 Oscar T. Grady es el gerente de producción de la compañía Citrus Groves, localizada justo al norte de Ocala, Florida. Oscar está preocupado debido a que las heladas tardías de los últimos tres años han estado dañando los 2,500 naranjos que posee la Citrus Groves. Con el fin de determinar el grado del daño ocasionado a los árboles, Oscar ha recogido una muestra del número de naranjas producidas por cada árbol para 42 naranjos y encontró que la producción promedio fue 525 naranjas por árbol, con una desviación estándar de 30 naranjas por árbol.

a) Estime la desviación estándar de la población a partir de la desviación estándar de la muestra.
b) Estime el error estándar de la muestra de esta población finita.
c) Construya un intervalo de confianza del 98% para la producción media por árbol del total de 2,500 árboles.
d) Si la producción media de naranjas por árbol fue 600 frutas hace cinco años, ¿qué puede decir Oscar acerca de la posible existencia de daños en el presente?

7-34 La jefa de policía, Kathy Ackert, recientemente estableció medidas enérgicas para combatir a los traficantes de droga de su ciudad. Desde que se pusieron en funcionamiento dichas medidas, han sido capturados 750 de los 12,368 traficantes de droga de la ciudad. El valor promedio, en dólares, de las drogas decomisadas a estos 750 traficantes es $250,000. La desviación estándar del valor de la droga de esos 750 traficantes es $41,000. Elabore para la jefa Ackert un intervalo de confianza del 90% para el valor medio en dólares de las drogas que están en manos de los traficantes de la ciudad.

Soluciones a los ejercicios de autoevaluación

EA 7-6 $\hat{\sigma} = 1.368 \quad N = 540 \quad n = 60 \quad \bar{x} = 6.2$

a) $\hat{\sigma}_{\bar{x}} = \dfrac{\hat{\sigma}}{\sqrt{n}} \times \sqrt{\dfrac{N-n}{N-1}} = \dfrac{1.368}{\sqrt{60}} \times \sqrt{\dfrac{540-60}{540-1}} = 0.167$

b) $\bar{x} \pm 2{,}05\hat{\sigma}_{\bar{x}} = 6.2 \pm 2.05(0.167) = 6.2 \pm 0.342 = (5.86, 6.54)$

EA 7-7 $s = 2.1 \quad n = 62 \quad \bar{x} = 24$

a) $\hat{\sigma} = s = 2.1$ psi

b) $\hat{\sigma}_{\bar{x}} = \hat{\sigma}/\sqrt{n} = 2.1/\sqrt{62} = 0.267$ psi

c) $\bar{x} \pm 1.96\hat{\sigma}_{\bar{x}} = 24 \pm 1.96(0.267) = 24 \pm 0.523 = (23.48, 24.52)$ psi

7.6 Cálculo de estimaciones de intervalo de la proporción a partir de muestras grandes

Repaso de la distribución binomial

Los especialistas en estadística, a menudo, utilizan una muestra para estimar la *proporción* de ocurrencias de un evento en una población. Por ejemplo, el gobierno estima, mediante un procedimiento de muestreo, el índice de desempleo o la proporción de personas sin trabajo de la fuerza laboral del país.

En el capítulo 5 introdujimos la distribución binomial, una distribución de datos discretos, no continuos. Presentamos, también, las dos fórmulas para derivar la media y la desviación estándar de la distribución binomial:

$$\mu = np \qquad [5\text{-}2]$$

$$\sigma = \sqrt{npq} \qquad [5\text{-}3]$$

donde,

- n = número de ensayos o intentos
- p = probabilidad de éxito
- $q = 1 - p$ = probabilidad de falla

Teóricamente, la distribución binomial es la distribución correcta a utilizar en la construcción de intervalos de confianza para estimar una proporción de población.

Limitaciones de la distribución binomial

Debido a que el cálculo de probabilidades binomiales es demasiado tedioso (recuerde que la probabilidad de obtener r éxitos en n ensayos es $[n!/r!(n-r)!][p^r q^{n-r}]$), el uso de la distribución binomial para elaborar estimaciones de intervalo de la proporción de una población es una proposición complicada. Afortunadamente, conforme aumenta el tamaño de la muestra, la distribución binomial puede aproximarse por una distribución normal apropiada, que podemos utilizar para aproximar la distribución muestral. Los estadísticos recomiendan que en la estimación, n sea lo suficientemente grande para que tanto np como nq sean al menos 5 cuando se utiliza la distribución normal como sustituto de la binomial.

Búsqueda de la media de la proporción de la muestra

Expresemos en símbolos la proporción de éxitos en una muestra con $\hat{p}$ (se lee *p gorro*). Luego modifiquemos la ecuación 5-2 de manera que podamos utilizarla para derivar la *media de la distribución de muestreo de la proporción de éxitos.* En palabras, $\mu = np$ muestra que la media de la distribución binomial es igual al producto del número de ensayos, n, por la probabilidad de obtener un éxito, p; esto es, np es igual al número medio de éxitos. Para cambiar este *número* de éxitos a la *proporción* de éxitos, dividimos np entre n y obtenemos sólo el valor de p. La media, que se encuentra al lado izquierdo de la ecuación se convierte en $\mu_{\hat{p}}$, es decir, en la media de la distribución de muestreo de la proporción de éxitos.

Media de la distribución muestral de la proporción	
$\mu_{\hat{p}} = p$	[7-3]

Búsqueda de la desviación estándar de la proporción de la muestra

De forma parecida podemos modificar la fórmula para la desviación estándar de la distribución binomial, $\sqrt{npq}$, que mide la desviación estándar del número de éxitos. Para cambiar el número de éxitos a la proporción de éxitos, dividimos $\sqrt{npq}$, entre n y obtenemos $\sqrt{pq/n}$. En términos estadísticos, la desviación estándar de la proporción de éxitos en una muestra se expresa en símbolos como:

Error estándar de la proporción

Error estándar de la proporción $\longrightarrow \sigma_{\hat{p}} = \sqrt{\dfrac{pq}{n}}$ [7-4]

y se conoce como el *error estándar de la proporción.*

Podemos ilustrar cómo utilizar estas fórmulas si, para una organización muy grande, hacemos la estimación de qué proporción de sus empleados prefieren planificar su propios beneficios de retiro en lugar de seguir un plan patrocinado por la compañía. Primero, tomamos una pequeña muestra aleatoria de 75 empleados y encontramos que el 0.4 de ellos están interesados en seguir sus propios planes de retiro. Nuestros resultados son:

$n = 75 \leftarrow$ Tamaño de muestra

$\hat{p} = 0.4 \leftarrow$ Proporción de la muestra a favor

$\hat{q} = 0.6 \leftarrow$ Proporción de la muestra en contra

Estimación de la proporción de una población

A continuación, la administración solicita que utilicemos esta muestra para encontrar un intervalo en el que puedan tener el 99% de confianza de que contiene a la proporción verdadera de la población.

Pero, para la *población,* ¿qué son $\hat{p}$ y $\hat{q}$? Podemos estimar los parámetros de la población mediante la sustitución de los estadísticos correspondientes de la muestra, $\hat{p}$ y $\hat{q}$ (*p gorro* y *q gorro*) en la fórmula del error estándar de la proporción.* Al hacer esto obtenemos:

Error estándar estimado de la proporción

Símbolo que indica que se está estimando el error estándar de la proporción

Estadístico de la muestra

$$\hat{\sigma}_{\hat{p}} = \sqrt{\frac{\hat{p}\hat{q}}{n}} \qquad [7\text{-}5]$$

$$= \sqrt{\frac{(0.4)(0.6)}{75}}$$

$$= \sqrt{0.0032}$$

$$= 0.057 \leftarrow \text{Error estándar estimado de la proporción}$$

Cálculo de los límites de confianza

Ahora estamos en posibilidades de proporcionar la estimación que la administración necesita, usando el mismo procedimiento que seguimos con anterioridad. Un nivel de confianza del 99% incluiría 49.5% del área que se encuentra a cualquier lado de la media de la distribución de muestreo. El cuerpo de la tabla 1 del apéndice nos dice que 0.495 del área bajo la curva normal está localizada entre la media y un punto que se encuentra a 2.58 errores estándar de la media. En consecuencia, 99% del área está contenida entre más *y* menos 2.58 errores estándar de la media. Nuestros límites de confianza entonces son:

* Note que no utilizamos el multiplicador de población finita, debido a que nuestra población es muy grande en comparación con el tamaño de la muestra.

$$\hat{p} + 2.58\,\hat{\sigma}_{\hat{p}} = 0.4 + 2.58(0.057)$$
$$= 0.4 + 0.147$$
$$= 0.547 \leftarrow \text{Límite superior de confianza}$$
$$\hat{p} - 2.58\,\hat{\sigma}_{\hat{p}} = 0.4 - 2.58(0.057)$$
$$= 0.4 - 0.147$$
$$= 0.253 \leftarrow \text{Límite inferior de confianza}$$

Nuestra conclusión

Entonces, estimamos a partir de nuestra muestra de 75 empleados que, con el 99% de confianza, creemos que la proporción de la población total de empleados que desean establecer sus propios planes de retiro está entre 0.253 y 0.547.

SUGERENCIAS Y SUPOSICIONES

Las mismas suposiciones, sugerencias y advertencias establecidas en la página 293 se aplican en este caso. La única diferencia es que ahora, como se trata de una proporción, la distribución binomial es la distribución muestral correcta. Recuerde, del capítulo 5, que mientras n sea suficientemente grande para que tanto np como nq sean al menos 5, se puede usar la distribución normal para aproximar la binomial. Si éste es el caso, se procede justo como se hizo con las estimaciones de intervalo de la media. Advertencia: como el error estándar exacto de la proporción depende de la proporción desconocida de la población (p), debe estimar p mediante $\hat{p}$, y usar $\hat{p}$, en la ecuación 7.5 para estimar el error estándar de la proporción.

Ejercicios 7.6

Ejercicios de autoevaluación

EA 7-8 Cuando se sondeó una muestra de 70 ejecutivos de ventas respecto al bajo desempeño durante noviembre en la industria de ventas al menudeo, el 66% pensó que la disminución en las ventas se debía a las temperaturas inusualmente altas, haciendo que los consumidores retrasaran sus compras de artículos de invierno.

a) Estime el error estándar de la proporción de ejecutivos de ventas que culpan al clima caliente de las bajas ventas.
b) Encuentre los límites de confianza superior e inferior para esta proporción dado un 95% de nivel de confianza.

EA 7-9 El doctor Benjamin Shockley, un psicólogo social reconocido, entrevistó a 150 ejecutivos de alto nivel y encontró que 42% de ellos no podía sumar fracciones correctamente.

a) Estime el error estándar de la proporción.
b) Construya un intervalo de confianza del 99% para la proporción verdadera de ejecutivos de alto nivel que no puede sumar fracciones correctamente.

Aplicaciones

■ **7-35** Pascal Inc., una tienda de computación que compra al mayoreo chips sin probar para computadora, está considerando cambiar a su proveedor por otro que se los ofrece probados y con garantía, a un precio más alto. Con el fin de determinar si éste es un plan costeable, Pascal debe determinar la proporción de chips defectuosos que le entrega el proveedor actual. Se probó una muestra de 200 chips y 5% tenía defectos.

a) Estime el error estándar de la proporción de chips defectuosos.
b) Construya un intervalo de confianza del 98% para la proporción de chips defectuosos adquiridos.

■ **7-36** General Cinema obtuvo una muestra de 55 personas que vieron *Caza Fantasmas 8* y les preguntaron si planeaban verla de nuevo. Sólo 10 de ellos pensaron que valía la pena ver la película por segunda vez.

a) Estime el error estándar de la proporción de asistentes al cine que verán la película por segunda vez.
b) Construya un intervalo de confianza del 90% para esta proporción.

■ **7-37** La encargada de publicidad para el nuevo postre garapiñado de lima-limón de los productos Clear´n Light está intranquila por el mal desempeño del postre en el mercado y por su futuro en la empresa. Preocupada porque su estrategia de comercialización no ha producido una identificación apropiada de las características del producto, tomó una muestra de 1,500 consumidores y encontró que 956 de éstos pensaban que el producto era una cera para pulir pisos.
   a) Estime el error estándar de la proporción de personas que tuvo esta grave interpretación errónea del postre.
   b) Construya un intervalo de confianza del 96% para la proporción verdadera de la población.

■ **7-38** Michael Gordon, un jugador profesional de básquetbol, lanzó 200 tiros de *castigo* y encestó 174 de ellos.
   a) Estime el error estándar de la proporción de todos los tiros que Michael falla.
   b) Construya un intervalo de confianza del 98% para la proporción de todos los tiros de *castigo* que Michael falla.

■ **7-39** Hace poco SnackMore encuestó a 95 consumidores y encontró que el 80% compraba galletas sin grasa de SnackMore cada mes.
   a) Estime el error estándar de la proporción.
   b) Construya un intervalo del 95% de confianza para la proporción verdadera de personas que compran las galletas cada mes.

■ **7-40** El dueño de la empresa Home Loan Company investigó aleatoriamente 150 de las 3,000 cuentas de la compañía y determinó que el 60% estaba en una posición excelente.
   a) Encuentre un intervalo de confianza del 95% para la proporción de cuentas que están en posición excelente.
   b) Con base en el inciso anterior, ¿qué tipo de estimación de intervalo podría dar para el número absoluto de cuentas que cumplen con el requisito de excelencia, manteniendo el mismo nivel de confianza del 95%?

■ **7-41** Durante un año y medio las ventas han estado disminuyendo de manera consistente en las 1,500 sucursales de una cadena de comida rápida. Una empresa de asesores ha determinado que el 31% de una muestra de 95 sucursales tiene claros signos de una mala administración. Construya un intervalo de confianza del 98% para esta proporción.

■ **7-42** El consejo estudiantil de una universidad tomó una muestra de 45 libros de texto de la librería universitaria y determinó que de ellos, 60% se vendía en más del 50% arriba de su costo al mayoreo. Dé un intervalo de confianza del 96% para la proporción de libros cuyo precio sea más del 50% mayor que el costo al mayoreo.

■ **7-43** Barry Turnbull, el famoso analista de Wall Street, está interesado en conocer la proporción de accionistas individuales que planean vender al menos un cuarto del total de sus valores el mes próximo. Barry ha efectuado una inspección aleatoria de 800 individuos que poseen acciones y ha establecido que el 25% de su muestra planea vender al menos la cuarta parte de sus acciones el mes siguiente. Barry está a punto de publicar su esperado informe mensual, "Pulso de Wall Street: indicador de cotizaciones", y le gustaría poder dar un intervalo de confianza a sus lectores. Está más preocupado por estar en lo correcto que por el ancho del intervalo. Construya un intervalo de confianza del 90% para la proporción verdadera de accionistas individuales que planean vender al menos un cuarto de sus acciones durante el siguiente mes.

Soluciones a los ejercicios de autoevaluación

EA 7-8 $n = 70 \qquad \hat{p} = 0.66$

a) $\hat{\sigma}_{\hat{p}} = \sqrt{\dfrac{\hat{p}\hat{q}}{n}} = \sqrt{\dfrac{0.66(0.34)}{70}} = 0.0566$

b) $\hat{p} \pm 1.96\hat{\sigma}_{\hat{p}} = 0.66 \pm 1.96(0.0566) = 0.66 \pm 0.111 = (0.549, 0.771)$

EA 7-9 $n = 150 \qquad \hat{p} = 0.42$

a) $\hat{\sigma}_{\hat{p}} = \sqrt{\dfrac{\hat{p}\hat{q}}{n}} = \sqrt{\dfrac{0.42(0.58)}{150}} = 0.0403$

b) $\hat{p} \pm 2.58\hat{\sigma}_{\hat{p}} = 0.42 \pm 2.58(0.0403) = 0.42 \pm 0.104 = (0.316, 0.524)$

7.7 Estimaciones de intervalos con la distribución *t*

En los tres ejemplos anteriores, los tamaños de la muestra eran todos mayores a 30. Muestreamos 100 limpiadores de parabrisas, 50 familias residentes de una área de cuatro manzanas de una comunidad y 75 empleados de una empresa grande. En cada ejemplo, la distribución normal era la distribución de muestreo adecuada para determinar intervalos de confianza.

A veces la distribución normal no es apropiada

Sin embargo, no siempre es éste el caso. ¿Cómo podríamos tratar estimaciones en las que la distribución normal *no* es la distribución de muestreo adecuada, es decir, cuando se estima la desviación estándar de la población y el tamaño de muestra es 30 o menos? Por ejemplo, en el problema con que abrimos el capítulo, referente al uso del carbón, tenemos datos que sólo comprenden 10 semanas. Afortunadamente, existe otra distribución que sí es apropiada para estos casos. Se conoce como *distribución t.*

Antecedentes históricos de la distribución *t*

Los primeros trabajos teóricos sobre la distribución *t* fueron realizados por W. S. Gosset, a principios del siglo XX. Gosset era empleado de la Cervecería Guinness en Dublín, Irlanda; la empresa no permitía que los empleados publicaran sus hallazgos de investigación con su propio nombre. De modo que Gosset adoptó el seudónimo de *Student* para publicar. En consecuencia, la distribución *t* se conoce como *distribución t de Student* o simplemente *distribución de Student*.

Condiciones para usar la distribución *t*

Debido a que se usa cuando el tamaño de la muestra es 30 o menos, los especialistas en estadística, suelen asociar la distribución *t* con estadísticas de muestras pequeñas. Esto es una mala interpretación porque el tamaño de la muestra es sólo *una* de las condiciones que nos llevan a utilizar la distribución *t*; la segunda es que la desviación estándar de la población debe ser desconocida. **El uso de la distribución *t* para hacer estimaciones se requiere siempre que el tamaño de la muestra sea menor o igual que 30 y la desviación estándar de la población no se conozca. Además, al utilizar la distribución *t*, suponemos que la población es normal o aproximadamente normal.**

Características de la distribución *t*

La distribución *t* comparada con la distribución normal

Sin derivar la distribución *t* de manera matemática, podemos entender en forma intuitiva la relación que existe entre la distribución *t* y la distribución *normal*. Ambas son simétricas. En general, la distribución *t* es más plana que la distribución normal y hay una distribución *t* diferente para cada tamaño posible de muestra. Aún así, conforme el tamaño de muestra se hace más grande, la forma de la distribución *t* deja de ser plana y se aproxima más a la distribución normal. De hecho, para tamaños de muestra mayores que 30, la distribución *t* se asemeja tanto a la normal que utilizaremos la normal para aproximar a la distribución *t*.

La figura 7-3 compara una distribución normal con dos distribuciones *t* para tamaños de muestra diferentes. En esta figura se muestran dos características de las distribuciones *t*. **Una distribución *t* es menor en la media y mayor en las colas que una distribución normal.** La figura también muestra cómo la distribución de Student tiene, proporcionalmente, una parte mayor de su área en las colas que la distribución normal; por esto será necesario alejarse más de la media de una distribución *t* para poder incluir la misma área bajo la curva. Entonces, los anchos de intervalo de una distribución de Student son mayores que los basados en la distribución normal.

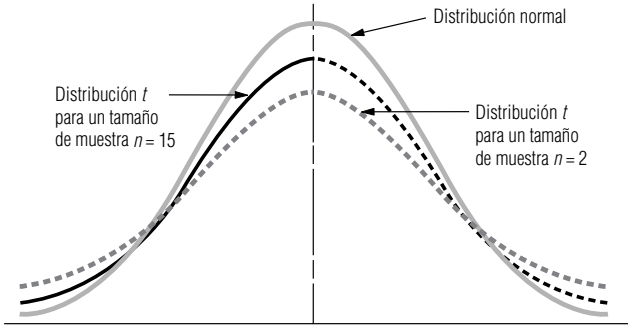
Grados de libertad

Definición de grados de libertad

Se afirmó que existe una distribución *t* diferente para cada tamaño de muestra. En un lenguaje estadístico apropiado, diríamos: "existe una distribución *t* distinta para cada uno de los *grados de libertad* posibles". **¿Qué son los grados de libertad? Podemos definirlos como el número de valores que podemos escoger libremente.**

FIGURA 7-3

Distribución normal y distribución *t* para una muestra $n = 15$, y distribución *t* para una muestra de tamaño $n = 2$



Suponga que se manejan dos valores de muestra, *a* y *b*, y sabemos que tienen una media de 18. En símbolos, la situación es:

$$\frac{a + b}{2} = 18$$

¿Cómo podemos encontrar los valores que *a* y *b* pueden tomar en esta situación? La respuesta es que *a* y *b* pueden ser cualesquiera dos valores cuya suma sea 36, ya que $36 \div 2 = 18$.

Suponga que sabemos que el valor de *a* es 10. Ahora *b* ya no es libre de tomar cualquier valor, sino que *debe* ser 26, ya que:

Si	$a = 10$
entonces	$\frac{10 + b}{2} = 18$
de modo que	$10 + b = 36$
por tanto	$b = 26$

Este ejemplo nos muestra que cuando hay dos elementos en una muestra y conocemos la media muestral de esos dos elementos, entonces somos libres de especificar sólo uno de los elementos, porque el otro estará determinado por el hecho de que los dos elementos suman el doble de la media de la muestra. En un lenguaje estadístico decimos que "tenemos un grado de libertad".

Otro ejemplo

Veamos otro ejemplo. Existen siete elementos en nuestra muestra y sabemos que la media de estos elementos es 16. En símbolos tenemos la siguiente situación:

$$\frac{a + b + c + d + e + f + g}{7} = 16$$

En este caso, los grados de libertad o el número de variables que podemos especificar libremente es $7 - 1 = 6$. Tenemos la libertad de asignar valores a seis variables, y luego ya no tenemos libertad de especificar el valor de la séptima variable; ésta queda determinada automáticamente.

Con dos valores de muestra tenemos un grado de libertad ($2 - 1 = 1$), y con siete valores de muestra tenemos seis grados de libertad ($7 - 1 = 6$). Entonces, en cada uno de estos dos ejemplos tenemos $n - 1$ grados de libertad, si *n* es el tamaño de la muestra. Similarmente, una muestra de 23 elementos nos daría 22 grados de libertad.

Función de los grados de libertad

Utilizaremos los grados de libertad cuando elijamos una distribución *t* para estimar una media de población, y utilizaremos $n - 1$ grados de libertad, cuando *n* es igual al tamaño de la muestra. Por ejemplo, si utilizamos una muestra de 20 para estimar una media de población, usaremos 19 grados de libertad para elegir la distribución *t* apropiada.

Uso de la tabla de distribución *t*

La tabla *t* comparada con la tabla *z*: tres diferencias

La tabla de los valores de la distribución *t* (tabla 2 del apéndice) difiere en su construcción de la tabla *z* que usamos antes. **La tabla *t* es más compacta y muestra áreas y valores de *t* sólo para algunos porcentajes (10, 5, 2 y 1%).** Debido a que hay una distribución *t* diferente para cada número de grados de libertad, una tabla más completa sería bastante grande. A pesar de que nos damos cuenta de la necesidad de una tabla más completa, de hecho la tabla 2 del apéndice contiene todos los valores de la distribución *t* que más se utilizan.

La segunda diferencia de la tabla *t* es que *no* se concentra en la probabilidad de que el parámetro de población que se está estimando se encuentre *dentro* del intervalo de confianza. En lugar de ello, mide la probabilidad de que el parámetro de población que estamos estimando *no* esté dentro de nuestro intervalo de confianza (es decir, la probabilidad de que esté *fuera*). Si estamos haciendo una estimación a un nivel de confianza del 90%, buscaríamos en la tabla *t* en la columna de 0.10 (100% − 90% = 10%). Esta probabilidad de 0.10 del error se representa con el símbolo α, la letra griega *alfa*. Encontraríamos los valores *t* apropiados para intervalos de confianza del 95, 98 y 99% en las columnas α con títulos 0.05, 0.02 y 0.01, respectivamente.

La tercera diferencia al utilizar la tabla *t* es que debemos especificar los grados de libertad que se manejan. Suponga que hacemos una estimación a un nivel de confianza del 90% con una muestra de tamaño 14, que tiene 13 grados de libertad. Busque en la tabla 2 del apéndice, en la columna de 0.10, hasta que encuentre el renglón 13. Del mismo modo que el valor *z*, el valor *t* de 1.771 indica que si señalamos una distancia de más menos $1.771\hat{\sigma}_{\bar{x}}$ (errores estándar estimados de $\bar{x}$) a ambos lados de la media, el área bajo la curva que se encuentra entre estos dos límites será el 90% del área total, y el área que se encuentra fuera de estos límites (la posibilidad de error) será el 10% del área total (vea la figura 7-4).

Recuerde que en el problema con que abrimos el capítulo, el administrador de la planta generadora de energía deseaba estimar la cantidad de carbón que requeriría este año, y tomó una muestra midiendo la cantidad de carbón utilizado durante 10 semanas. Los datos de la muestra son:

$n = 10$ semanas ← Tamaño de la muestra

$gl = 9$ ← Grados de libertad

$\bar{x} = 11{,}400$ toneladas ← Media de la muestra

$s = 700$ toneladas ← Desviación estándar de la muestra

Uso de la tabla *t* para calcular límites de confianza

El administrador de la planta desea una estimación de intervalo del consumo medio de carbón, y quiere estar 95% seguro de que el consumo medio se encuentre dentro de dicho intervalo. **Este problema requiere el uso de una distribución *t*, porque el tamaño de la muestra es menor que 30, no se conoce la desviación estándar de la población y el administrador piensa que la población es aproximadamente normal.**

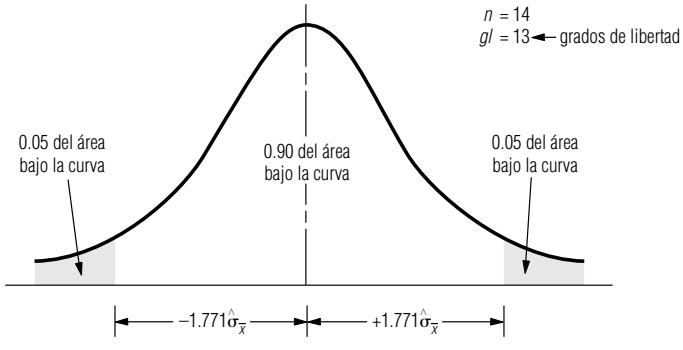


FIGURA 7-4

Distribución *t* para 13 grados de libertad que muestra un intervalo de confianza del 90%

Como primer paso para resolver este problema, recuerde que *estimamos* la desviación estándar de la población a partir de la desviación estándar de la muestra; por consiguiente:

$$\hat{\sigma} = s \qquad [7\text{-}1]$$

$$= 700 \text{ toneladas}$$

Con esta estimación de la desviación estándar de la población, podemos estimar el error estándar de la media si modificamos la ecuación 7-2 para omitir el multiplicador de población finita (debido a que el tamaño de muestra de 10 semanas es menor que el 5% de cinco años, 260 semanas, periodo para el que se tienen datos disponibles):

Error estándar estimado de la media de una población infinita

$$\hat{\sigma}_{\bar{x}} = \frac{\hat{\sigma}}{\sqrt{n}} \qquad [7\text{-}6]$$

Prosiguiendo con nuestro ejemplo, encontramos que $\hat{\sigma}_{\bar{x}} = \dfrac{700}{\sqrt{10}}$

$$= \frac{700}{3.162}$$

$$= 221.38 \text{ toneladas} \leftarrow \text{Error estándar estimado de la media de una población infinita}$$

Ahora buscamos en la tabla 2 del apéndice en la columna 0.05 (100% − 95% = 5%) y el renglón de 9 grados de libertad (10 − 1 = 9). Vemos que el valor *t* es 2.262 y con él podemos establecer nuestros límites de confianza:

$$\bar{x} + 2.262\hat{\sigma}_{\bar{x}} = 11{,}400 \text{ toneladas} + 2.262(221.38 \text{ toneladas})$$

$$= 11{,}400 + 500.76$$

$$= 11{,}901 \text{ toneladas} \leftarrow \text{Límite superior de confianza}$$

$$\bar{x} - 2.262\hat{\sigma}_{\bar{x}} = 11{,}400 \text{ toneladas} - 2.262(221.38 \text{ toneladas})$$

$$= 11{,}400 - 500.76$$

$$= 10{,}899 \text{ toneladas} \leftarrow \text{Límite inferior de confianza}$$

Nuestra conclusión

El intervalo de confianza se ilustra en la figura 7-5. Ahora podemos informar al administrador de la planta con el 95% de confianza que el consumo medio semanal de carbón se encuentra entre 10,899 y 11,901 toneladas, y el administrador puede utilizar la cifra de 11,901 toneladas para estimar la cantidad de carbón a ordenar.

La única diferencia entre el proceso utilizado para hacer esta estimación y los procedimientos para resolver los problemas anteriores es el uso de la distribución *t* como la distribución adecuada. **Recuerde que en cualquier problema de estimación donde el tamaño de la muestra sea menor o igual que 30, la desviación estándar de la población no se conozca *y* la población en cuestión sea normal o aproximadamente normal, utilizamos la distribución *t*.**

Resumen de los límites de confianza en condiciones diferentes

En la tabla 7-5 resumimos los diferentes planteamientos para la estimación introducidos en este capítulo y los límites de confianza apropiados para cada uno.

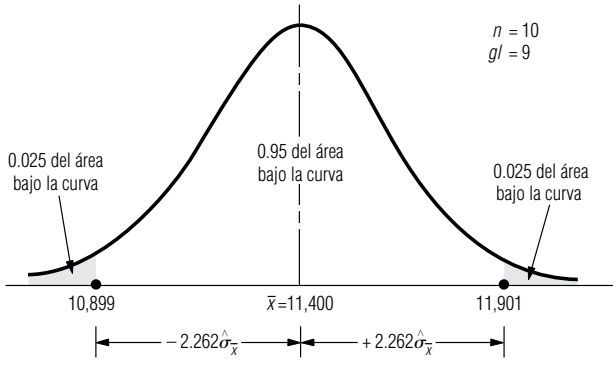


FIGURA 7-5

Problema del carbón: distribución *t* con 9 grados de libertad y un intervalo de confianza del 95%

Tabla 7-5

Resumen de las fórmulas para límites de confianza en la estimación de la media y la proporción

	Cuando la población es finita (y $n/N > 0.05$)	Cuando la población es infinita (o $n/N < 0.05$)
Estimación de μ (la media de la población): Cuando σ (la desviación estándar de la población) se conoce	Límite superior: $\bar{x} + z\frac{\sigma}{\sqrt{n}} \times \sqrt{\frac{N-n}{N-1}}$	$\bar{x} + z\frac{\sigma}{\sqrt{n}}$
	Límite inferior: $\bar{x} - z\frac{\sigma}{\sqrt{n}} \times \sqrt{\frac{N-n}{N-1}}$	$\bar{x} + z\frac{\sigma}{\sqrt{n}}$
Cuando σ (la desviación estándar de la población) no se conoce ($\hat{\sigma} = s$) Cuando n (el tamaño de la muestra) es mayor que 30	Límite superior: $\bar{x} + z\frac{\hat{\sigma}}{\sqrt{n}} \times \sqrt{\frac{N-n}{N-1}}$	$\bar{x} + z\frac{\hat{\sigma}}{\sqrt{n}}$
	Límite inferior: $\bar{x} - z\frac{\hat{\sigma}}{\sqrt{n}} \times \sqrt{\frac{N-n}{N-1}}$	$\bar{x} - z\frac{\hat{\sigma}}{\sqrt{n}}$
Cuando n (el tamaño de la muestra) es 30 o menos y la población es normal o aproximadamente es normal*	Este caso está más allá del objetivo del libro; consulte a un especialista en estadística.	$\bar{x} + t\frac{\hat{\sigma}}{\sqrt{n}}$
		$\bar{x} - t\frac{\hat{\sigma}}{\sqrt{n}}$
Estimación de p (la proporción de la población): Cuando n (el tamaño de la muestra) es mayor que 30	Este caso está más allá del objetivo del libro; consulte a un especialista en estadística.	$\hat{p} + z\hat{\sigma}_{\hat{p}}$
$\hat{\sigma}_{\hat{p}} = \sqrt{\frac{\hat{p}\hat{q}}{n}}$		$\hat{p} - z\hat{\sigma}_{\hat{p}}$

*Recuerde que la distribución *t* apropiada es la que tiene $n - 1$ grados de libertad.

SUGERENCIAS Y SUPOSICIONES

El concepto de *grados de libertad* suele ser difícil de entender al principio. Sugerencia: piense en los grados de libertad como el número de opciones con que cuenta. Si hay mantequilla de maní y queso en el refrigerador, se puede elegir un emparedado de mantequilla de maní o uno de queso (a menos que le gusten los emparedados de mantequilla de maní con queso). Si al abrir la puerta ve que ya no hay queso, el señor Gosset tal vez diría, "ahora tiene cero grados de libertad". Esto es, si desea almorzar, no tiene opciones; come mantequilla de maní o muere de hambre. Advertencia: aunque la distribución *t* está asociada con las estadísticas de muestras pequeñas, recuerde que un tamaño de muestra menor que 30 es sólo una de las condiciones para usarla. Las otras son que no se conozca la desviación estándar de la población y que la población siga una distribución normal o una aproximadamente normal.

Ejercicios 7.7

Ejercicios de autoevaluación

EA 7-10 Para los siguientes tamaños de muestra y niveles de confianza, encuentre los valores *t* adecuados para construir intervalos de confianza:

a) $n = 28$; 95%.
b) $n = 8$; 98%.
c) $n = 13$; 90%.
d) $n = 10$; 95%.
e) $n = 25$; 99%.
f) $n = 10$; 99%.

EA 7-11 Se obtuvo una muestra aleatoria de siete amas de casa y se determinó que las distancias caminadas al realizar las tareas domésticas dentro de la casa tenían un promedio de 39.2 millas por semana y una desviación estándar de la muestra de 3.2 millas por semana. Construya un intervalo de confianza del 95% para la media de la población.

Conceptos básicos

■ **7-44** Para los siguientes tamaños de muestra y niveles de confianza, encuentre los valores *t* adecuados para construir intervalos de confianza:

a) $n = 15$; 90%.
b) $n = 6$; 95%.
c) $n = 19$; 99%.
d) $n = 25$; 98%.
e) $n = 10$; 99%.
f) $n = 41$; 90%.

■ **7-45** Dados los siguientes tamaños de muestra y los valores *t* utilizados para construir intervalos de confianza, encuentre los niveles de confianza correspondientes:

a) $n = 27$; $t = \pm 2.056$.
b) $n = 5$; $t = \pm 2.132$.
c) $n = 18$; $t = \pm 2.898$.

■ **7-46** Una muestra de 12 elementos tiene una media de 62 y una desviación estándar de 10. Construya un intervalo de confianza del 95% para la media de la población.

■ **7-47** La siguiente muestra de ocho observaciones fue tomada de una población infinita con distribución normal:

75.3 76.4 83.2 91.0 80.1 77.5 84.8 81.0

a) Encuentre la media.
b) Estime la desviación estándar de la población.
c) Construya un intervalo de confianza del 98% para la media.

Aplicaciones

■ **7-48** Las autoridades de la parte norte del condado de Orange han encontrado, para consternación de los comisionados del condado, que la población presenta severos problemas relacionados con placa dentobacteriana. Cada año, el departamento de salud dental local examina una muestra tomada de los habitantes del condado y registra la condición de la dentadura de cada paciente en una escala de 1 a 100, donde 1 indica que no hay placa dentobacteriana y 100 indica que es muy grande. Este año, el departamento de salud dental examinó a 21 pacientes y encontró que tenían un promedio de placa dentobacteriana de 72 con una desviación estándar de 6.2. Construya un intervalo de confianza del 98% para la media del índice de placa dentobacteriana de la parte norte de Orange.

■ **7-49** Se obtuvo una muestra aleatoria de 12 cajeros de banco y se determinó que cometían un promedio de 3.6 errores por día con una desviación estándar muestral de 0.42 errores. Construya un intervalo del 90% de confianza para la media de la población de errores por día. ¿Qué suposición está implícita acerca del número de errores que cometen los cajeros?

■ **7-50** La senadora Hanna Rowe ha ordenado que se haga una investigación acerca del gran número de accidentes en bote que han ocurrido en el estado durante los últimos veranos. Siguiendo sus instrucciones, su ayudante, Geoff Spencer, ha seleccionado al azar 9 meses de verano entre los últimos años y ha recabado datos acerca de los accidentes en bote ocurridos en cada uno de esos meses. El número medio de accidentes que se presentaron en los 9 meses fue 31, y la desviación estándar de esta muestra fue 9 accidentes por mes. Se pidió a Geoff que construyera un intervalo de confianza del 90% para el número real de accidentes por mes, pero él mismo sufrió un accidente en bote recientemente, por lo que usted tendrá que terminar su trabajo.

Soluciones a los ejercicios de autoevaluación

EA 7-10 a) 2.052.
b) 2.998.
c) 1.782.
d) 2.262.
e) 2.797.
f) 3.250.

EA 7-11 $s = 3.2 \quad n = 7 \quad \bar{x} = 39.2 \quad \hat{\sigma}_{\bar{x}} = s/\sqrt{n} = 3.2/\sqrt{7} = 1.2095$

$$\bar{x} \pm t\hat{\sigma}_{\bar{x}} = 39.2 \pm 2.447(1.2095) = 39.2 \pm 2.9596$$
$$= (36.240, 42.160) \text{ millas}$$

7.8 Determinación del tamaño de muestra en estimación

En todos los análisis hechos hasta ahora, hemos utilizado el símbolo *n* en lugar de un número específico. Ahora necesitamos saber cómo determinar el número que se debe usar. ¿Qué tan grande deberá ser la muestra? Si ésta es muy pequeña, podemos fallar en el logro de los objetivos de nuestro análisis; si es demasiado grande, desperdiciamos recursos al tomar la muestra.

¿Cuál es el tamaño adecuado de la muestra?

Se presentará cierto grado de error de muestreo por no estudiar a la población completa. Siempre que tomamos una muestra, perdemos *algo* de información útil de la población. Si queremos tener un alto nivel de precisión (esto es, si deseamos estar bastante seguros de nuestra estimación), debemos muestrear la población lo suficiente para asegurarnos que obtuvimos la información requerida. El error de muestreo se puede controlar si seleccionamos una muestra con el tamaño adecuado. En general, cuanta más precisión se quiera, más grande será el tamaño necesario de la muestra. Examinemos algunos métodos útiles en la determinación del tamaño necesario de muestra para cualquier nivel específico de precisión.

Tabla 7-6	Límite inferior de confianza	Límite superior de confianza
Comparación de dos maneras de expresar los mismos límites de confianza	a. $\bar{x} - \$500$ b. $\bar{x} - z\sigma_{\bar{x}}$	a. $\bar{x} + \$500$ b. $\bar{x} + z\sigma_{\bar{x}}$

Tamaño de muestra para estimar una media

Suponga que una universidad está efectuando una investigación acerca de los ingresos anuales de los estudiantes del último año de su escuela de administración. Se sabe, por experiencia, que la desviación estándar de los ingresos anuales de la población completa (1,000 estudiantes) de los egresados es alrededor de $1,500. ¿Qué tan grande debe ser la muestra que debe tomar la universidad con el fin de estimar el ingreso medio anual de los estudiantes graduados el año pasado, dentro de más menos $500 y con un nivel de confianza del 95%?

¿Exactamente qué se pide en este problema? La universidad va a tomar una muestra de *cierto* tamaño, determinará la media de la muestra, $\bar{x}$, y la usará como estimación puntual de la media de la población. Quiere tener la certeza del 95% de que el ingreso medio anual real de la generación de graduados el año pasado no esté más de $500 arriba o abajo de la estimación puntual. El renglón *a* de la tabla 7-6 resume, en símbolos, la forma en que la universidad define sus límites de confianza. En el renglón *b* se muestran los símbolos para expresar los límites de confianza para una población infinita. Cuando comparamos estos dos conjuntos de límites de confianza, podemos ver que:

Dos maneras de expresar un límite de confianza

$$z\sigma_{\bar{x}} = \$500$$

Así, la directiva de la universidad en realidad está diciendo que desea que $z\sigma_{\bar{x}}$ sea igual a $500. Si buscamos en la tabla 1 del apéndice el valor necesario de z para un nivel de confianza del 95%, vemos que es 1.96. Paso a paso:

$$\begin{aligned}
&\text{si} && z\sigma_{\bar{x}} = \$500\\
&\text{y} && z = 1.96\\
&\text{entonces} && 1.96\sigma_{\bar{x}} = \$500\\
&\text{y} && \sigma_{\bar{x}} = \frac{\$500}{1.96}\\
& && = \$255 \leftarrow \text{Error estándar de la media}
\end{aligned}$$

Recuerde que la fórmula para el error estándar es la ecuación 6-1:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \leftarrow \text{Desviación estándar de la población} \qquad [6\text{-}1]$$

Búsqueda de un tamaño de muestra adecuado

Utilizando la ecuación 6-1, podemos sustituir el valor conocido de la desviación estándar de la población, $1,500, y el valor calculado del error estándar de $255 y despejar n:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \qquad [6\text{-}1]$$

$$\$255 = \frac{\$1{,}500}{\sqrt{n}}$$

$$(\sqrt{n})(\$255) = \$1{,}500$$

$$\sqrt{n} = \frac{\$1{,}500}{\$255}$$

$$\sqrt{n} = 5.882 \text{; ahora elevamos al cuadrado ambos lados}$$

$$n = 34.6 \leftarrow \text{Tamaño de muestra para la precisión especificada}$$

Por tanto, como n debe ser mayor o igual que 34.6, la universidad deberá tomar una muestra de 35 graduados el año pasado de la escuela de administración para obtener la precisión que desea en la estimación del ingreso medio anual de la generación.

Estimación de la desviación estándar a partir del rango

En el ejemplo anterior conocíamos la desviación estándar de la población, pero en muchos otros casos no está disponible. Recuerde, también, que todavía no hemos tomado la muestra y que estamos intentando decidir de qué tamaño va a ser. No podemos estimar la desviación estándar de la población utilizando los métodos presentados en la primera parte del capítulo. Pero si tenemos idea de cuál es el rango de la población, podemos utilizarlo para obtener una estimación burda pero manejable de la desviación estándar.

Suponga que estamos estimando el índice de salarios de manufactura por hora en una ciudad, y que tenemos bastante seguridad de que existe una diferencia de $4.00 entre el índice más alto y el más bajo. Sabemos que más y menos 3 desviaciones estándar incluyen el 99.7% del área total bajo la curva normal, esto es, más 3 desviaciones estándar y menos 3 desviaciones estándar de la media incluyen a casi toda el área de la distribución. Para representar esta relación, hemos construido la figura 7-6, en la cual $4.00 (el rango) es igual a 6 desviaciones estándar (más 3 y menos 3). Por consiguiente, una estimación burda de la desviación estándar de la población sería:

$$6\hat{\sigma} = \$4.00$$

$$\hat{\sigma} = \frac{\$4.00}{6}$$

$$\text{Estimación de la desviación estándar de lo población} \rightarrow \hat{\sigma} = \$0.667$$

La estimación de la desviación estándar de la población obtenida con este método burdo, no es una estimación precisa, pero puede significar la diferencia entre obtener una idea que funcione del tamaño requerido de la muestra y no saber nada con respecto a ese tamaño de muestra.

Tamaño de muestra para estimar una proporción

Los procedimientos utilizados para determinar los tamaños de muestra para estimar una proporción de la población son parecidos a los que se utilizan para estimar una media de población. Suponga que deseamos encuestar a estudiantes de una universidad grande. Deseamos determinar qué proporción de éstos está a favor de un nuevo sistema de evaluación. Nos gustaría contar con un tamaño de muestra que nos permita tener una certeza del 90% de que estamos estimando la proporción verdadera de la población de 40,000 estudiantes a favor del nuevo sistema de evaluación, más menos 0.02.

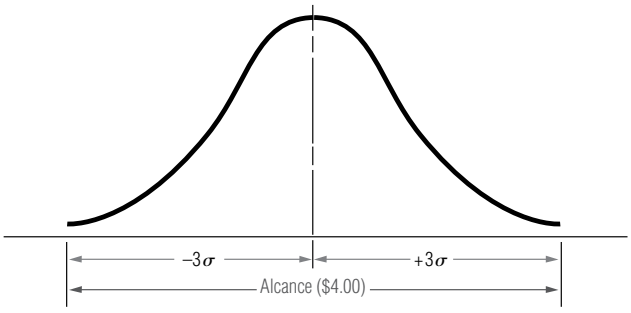


FIGURA 7-6

Relación aproximada entre el rango y la desviación estándar de la población

Empezamos a resolver este problema buscando en la tabla 1 del apéndice un valor de z correspondiente a un nivel de confianza del 90%. Tal valor es ± 1.64 errores estándar a partir de la media. Queremos que nuestra estimación esté dentro de 0.02, de modo que podemos simbolizar el proceso paso a paso de la siguiente manera:

$$\begin{aligned} &\text{Si} && z\sigma_{\hat{p}} = 0.02 \\ &\text{y} && z = 1.64 \\ &\text{entonces} && 1.64\sigma_{\hat{p}} = 0.02 \end{aligned}$$

Si ahora sustituimos los valores que se tienen para $\sigma_{\hat{p}}$ en la parte derecha de la ecuación 7-4, obtenemos:

$$1.64\sqrt{\frac{pq}{n}} = 0.02$$

$$\sqrt{\frac{pq}{n}} = 0.0122; \text{ ahora elevamos al cuadrado ambos lados}$$

$$\frac{pq}{n} = 0.00014884; \text{ ahora multiplicamos ambos lados por } n$$

$$pq = 0.00014884n$$

$$n = \frac{pq}{0.00014884}$$

Para hallar n, todavía necesitamos una estimación de los parámetros p y q de la población. Si tenemos una buena idea de la proporción real de estudiantes que están a favor del nuevo sistema, podemos utilizarla como nuestra mejor estimación para calcular n. Pero si no tenemos idea del valor de p, entonces nuestra mejor estrategia es darle un valor de manera tal que escogemos n en forma conservadora (es decir, de modo que el tamaño de la muestra *sea* lo suficientemente grande para darnos, al menos, la precisión que necesitamos sin importar el verdadero valor de p). En este punto del problema, n es igual al producto de p y q dividido entre 0.00014884. La manera de obtener la n más grande es generando el numerador más grande posible de esa expresión, lo cual sucede cuando elegimos $p = 0.5$ y $q = 0.5$. Entonces n se convierte en:

$$n = \frac{pq}{0.00014884}$$

$$= \frac{(0.5)(0.5)}{0.00014884}$$

$$= \frac{0.25}{0.00014884}$$

$$= 1{,}680 \leftarrow \text{Tamaño de muestra para la precisión especificada}$$

Como respuesta, para tener una seguridad del 90% de que estimamos la proporción verdadera dentro de 0.02, debemos escoger una muestra aleatoria simple de 1,680 estudiantes para entrevistar.

Selección de la proporción más conservadora

En el problema que acabamos de resolver, hemos tomado un valor para p que representó la estrategia más conservadora; el valor de 0.5 generó la muestra más grande posible. Habríamos utilizado otro valor de p si hubiéramos podido estimar uno o si hubiésemos tenido una buena idea de su valor. Siempre que estas dos últimas soluciones estén ausentes, tome el valor más conservador posible de p, a saber, $p = 0.5$.

Tabla 7-7
Tamaño de muestra n asociado con diferentes valores de p y q

Escoja este valor para p	Valor de q o $1 - p$	$\left(\frac{pq}{0.00014884}\right)$	Tamaño de muestra
0.2	0.8	$\frac{(0.2)(0.8)}{(0.00014884)}$	= 1,075
0.3	0.7	$\frac{(0.3)(0.7)}{(0.00014884)}$	= 1,411
0.4	0.6	$\frac{(0.4)(0.6)}{(0.00014884)}$	= 1,613
0.5	0.5	$\frac{(0.5)(0.5)}{(0.00014884)}$	= 1 ,680 ← El más conservador
0.6	0.4	$\frac{(0.6)(0.4)}{(0.00014884)}$	= 1,613
0.7	0.3	$\frac{(0.7)(0.3)}{(0.00014884)}$	= 1,411
0.8	0.2	$\frac{(0.8)(0.2)}{(0.00014884)}$	= 1,075

Para ilustrar que 0.5 produce el valor más grande posible para el tamaño de la muestra, en la tabla 7-7 se resuelve el problema del sistema de evaluación utilizando varios valores de p. Del tamaño de las muestras asociado con esos valores, puede ver que para el intervalo de valores de p que va de 0.3 a 0.7, el cambio en el tamaño de muestra correspondiente es relativamente pequeño. Por tanto, aunque ya hubiera sabido que la proporción de población verdadera es 0.3 y de todos modos usara 0.5, hubiera muestreado solamente 269 personas más (1,680 − 1,411) de lo que era realmente necesario para el grado de precisión deseado. Obviamente, adivinar valores de p en casos como éste no es tan crítico como parecía a primera vista.

SUGERENCIAS Y SUPOSICIONES

Desde una perspectiva de sentido común, si la desviación estándar de una población es muy pequeña, los valores se agrupan muy cerca de la media y casi cualquier tamaño de muestra los captará y producirá información precisa. Por otro lado, si la desviación estándar de la población es muy grande y los valores están bastante dispersos, será necesaria una muestra muy grande para incluirlos y obtener información correcta. ¿Cómo puede tenerse una idea de la desviación estándar de la población antes de iniciar el muestreo? Las compañías que planean realizar estudios de mercado casi siempre hacen una investigación preliminar de la población para estimar la desviación estándar. Si el producto se parece a otro que ha estado en el mercado, a menudo es posible apoyarse en los datos anteriores acerca de la población sin más estimaciones.

Ejercicios 7.8

Ejercicios de autoevaluación

EA 7-12 Para un mercado de prueba, encuentre el tamaño de la muestra requerido para estimar la proporción verdadera de consumidores satisfechos con cierto producto dentro de ± 0.04 en un nivel de confianza del 90%. Suponga que no se tiene una idea buena acerca de cuál es la proporción.

EA 7-13 Un curso de lectura rápida garantiza cierto aumento en la velocidad de lectura en 2 días. El profesor sabe que algunas personas no podrán lograr este incremento, de manera que antes de establecer el porcentaje garantizado de personas que lograrán el incremento en la velocidad de lectura, desea tener una confianza del 98% de que el porcentaje se ha estimado dentro de $\pm 5\%$ del valor verdadero. ¿Cuál es el tamaño de muestra más conservador necesario en este problema?

Conceptos básicos

- **7-51** Si la desviación estándar de la población es 78, encuentre el tamaño de muestra necesario para estimar la media verdadera dentro de 50 puntos, para un nivel de confianza del 95%.
- **7-52** Se tienen fuertes indicios de que la proporción es alrededor de 0.7. Encuentre el tamaño de muestra necesario para estimar la proporción dentro de ± 0.02 con un nivel de confianza del 90%.
- **7-53** Dada una población con una desviación estándar de 8.6, ¿qué tamaño de muestra es necesario para estimar la media de la población dentro de ± 0.5 con un nivel de confianza del 99%?

Aplicaciones

- **7-54** Debe votarse una propuesta importante y un político desea encontrar la proporción de personas que están a favor de la propuesta. Encuentre el tamaño de muestra requerido para estimar la proporción verdadera dentro de ± 0.05 con un nivel de confianza del 95%. Suponga que no se tiene idea de cuál es la proporción. ¿Cuál sería el cambio en el tamaño de la muestra si pensara que cerca del 75% de las personas favorece la propuesta? ¿Cuál sería el cambio si sólo alrededor del 25% favorece la propuesta?
- **7-55** La administración de la empresa Southern Textiles, recientemente ha sido atacada por la prensa debido a los supuestos efectos de deterioro en la salud que ocasiona su proceso de fabricación. Un sociólogo ha aventurado la teoría de que los empleados que mueren por causas naturales muestran una marcada consistencia en la duración de su vida: los límites superior e inferior de la duración de sus vidas no difieren en más de 550 semanas (alrededor de 10 $^1/_2$ años). Para un nivel de confianza del 98%, ¿qué tan grande debe ser la muestra, dentro de ± 30 semanas, que ha de examinarse para encontrar la vida promedio de estos empleados dentro de ± 30 semanas?
- **7-56** Food Tiger, una tienda local, vende bolsas de plástico para basura y ha recibido unas cuantas quejas respecto a su resistencia. Parece que las bolsas que vende son menos resistentes que las de su competidor y, en consecuencia, se rompen más a menudo. John C. Tiger, gerente de adquisiciones, está interesado en determinar el peso máximo promedio que puede resistir las bolsas para basura sin que se rompan. Si la desviación estándar del peso límite que rompe una bolsa es 1.2 kg, determine el número de bolsas que deben ser probadas con el fin de que el señor Tiger tenga una certeza del 95% de que el peso límite promedio está dentro de 0.5 kg del promedio verdadero.
- **7-57** La universidad está considerando la posibilidad de elevar la colegiatura con el fin de mejorar las instalaciones; para ello, sus autoridades desean determinar qué porcentaje de estudiantes están a favor del aumento. La universidad necesita tener una confianza del 90% de que el porcentaje se determinó dentro del 2% del valor verdadero. ¿Qué tamaño de muestra se requiere para garantizar esta precisión independientemente del porcentaje verdadero?
- **7-58** Wicks y Ticks, una tienda local especializada en velas y relojes está interesada en obtener una estimación de intervalo para el número medio de clientes que entran a la tienda diariamente. Los dueños tienen una seguridad razonable de que la desviación estándar real del número diario de clientes es 15. Ayude a Wicks y Ticks a salir de un bache determinando el tamaño de muestra que deberán utilizar para desarrollar un intervalo de confianza del 96% para la media verdadera que tenga un ancho de sólo ocho clientes.

Soluciones a los ejercicios de autoevaluación

EA 7-12 Suponga que $p = q = 0.5$.

$$0.04 = 1.64\sqrt{\frac{pq}{n}} = 1.64\sqrt{\frac{0.5(0.5)}{n}} \text{ así } n = \left(\frac{1.64(0.5)}{0.04}\right)^2 = 420.25 \text{ es decir, } n \geq 421.$$

EA 7-13 Suponga que $p = q = 0.5$.

$$0.05 = 2.33\sqrt{\frac{pq}{n}} = 2.33\sqrt{\frac{0.5(0.5)}{n}} \text{ así } n = \left(\frac{2.33(0.5)}{0.05}\right)^2 = 542.89 \text{ es decir, } n \geq 543.$$

Por tanto, debe tomarse una muestra de al menos 543 registros de estudiantes anteriores.

Estadística en el trabajo

Loveland Computers

Caso 7: Estimación Aunque Lee Azko se ha sentido un tanto nervioso en su primer trabajo, las tareas que se le han encomendado en producción y adquisiciones le han mostrado cómo aplicar lo que aprendió en los libros. El siguiente trabajo introdujo a Lee en otro departamento de Loveland Computers y lo enfrentó con el enfoque sin sentido de su directora, Margot Derby.

"Déjame explicarte la situación", comenzó Margot, dejando de lado cualquier preámbulo. "Ya sabes que nos consideramos, principalmente, distribuidores de equipos de cómputo: computadoras personales que la gente utiliza en sus negocios y casas. Cuando empezamos, dejamos que el cliente buscara el software. En algunas ocasiones, compran sus programas a las compañías que los diseñan o a distribuidores nacionales que atienden pedidos por teléfono. Ahora ya hay algunos distribuidores al menudeo locales; casi todos los centros comerciales suburbanos tienen al menos una tienda que vende programas de computación.

"La razón por la cual no vendemos software es que ya había demasiados programas en el mercado, y no queríamos adivinar cuál de ellos iba a ser el producto de mayor venta, equivocarnos y terminar con un inventario de programas inútiles. Pero la situación ha cambiado. Después de algunas sacudidas en el mercado del software, han surgido dos o tres líderes notables en cada campo; por ejemplo, hojas de cálculo y procesadores de palabras. Para equilibrar la competencia, empezamos a incluir algo de software en nuestras computadoras con fines de promoción.

"El año pasado, empezamos a cargar los programas en el disco duro para ciertos clientes. Podemos darles precios bastante competitivos por el software, y los programas precargados se convirtieron en una característica importante que mucha gente busca en el producto. Con estos antecedentes, estoy considerando nuevamente el software para ver si cambiamos nuestra estrategia y hacemos algo más en esa línea. Para darme una idea del mercado, pedí que interrogaran a 500 clientes que tienen una computadora Loveland desde hace aproximadamente un año; les preguntaron cuánto gastaron, en total, en software durante el primer año.

"Tengo todos los datos aquí; no me llevó ni dos minutos calcular la media y la desviación estándar con nuestro programa de hoja de cálculo. Los banqueros inversionistas de Nueva York le echaron una mirada a un borrador de mi plan de comercialización de software y, cuando vinieron la semana pasada, me preguntaron qué tan segura podía estar de que los resultados de mi investigación telefónica eran exactos.

"Cada vez que tomo el periódico, veo alguna encuesta de opinión en la que se dice algo como 'esta encuesta está basada en un sondeo de 1,200 adultos y tiene un margen de error del 3%'. ¿Cómo es que saben eso? ¿Tienen registros de todos los investigados y de cuándo están en lo correcto o no? Sólo tengo este conjunto de resultados y no veo cómo responder a las preguntas de los inversionistas."

"No debe ser muy difícil", respondió Lee al tiempo que inspeccionaba el escritorio para asegurarse de que había a la mano una calculadora y un conjunto de tablas estadísticas. "¿Por qué no me muestras los datos que tienes? Tal vez podamos darnos una idea de la respuesta ahora mismo."

Preguntas de estudio: ¿Qué distribución supondrá Lee que tienen los resultados de la encuesta telefónica, y qué tabla estadística será más útil? ¿Cómo puede Lee definir *margen de error* para Margot? ¿Es probable que Lee recomiende una muestra más grande?

Ejercicio de base de datos computacional

HH Industries

Al inicio de la siguiente semana, Bob regresó a la oficina de Laurel. "Bueno, hemos empezado a encuestar a nuestra muestra", comentó. "¿Podrías ayudarme a tener una idea de cuántos debemos examinar? Estoy interesado en un nivel de confianza del 95% de estar dentro de más menos 0.05 de la proporción verdadera de la población. Pienso que vas a estar de acuerdo conmigo en que, para fines prácticos, podemos considerar a nuestra población como infinita."

"Creo que tienes razón", acordó Laurel. "¡Ya vi la fila de archiveros! Para estimar el número a encuestar, sería de gran ayuda que tuviéramos una idea fundada del parámetro real de la población, pero al menos podremos obtener un intervalo de tamaños de muestra."

I. Determine un tamaño de muestra apropiado para satisfacer las condiciones de Bob, si el valor real de p (la proporción de órdenes de compra hechas de manera competitiva) es aproximadamente 0.2, 0.3, 0.4 o 0.5. ¿Cuál deberá escoger Bob?

Aproximadamente una semana después, Bob tocó en la puerta de Laurel. "Aquí están los datos sin procesar. El objetivo de Hal, en este punto, es que tengamos al menos el 60%

de las órdenes de compra hechas de manera competitiva. ¿Crees que esto lo pondrá contento?"

"Calculemos nuestro intervalo de confianza y ya veremos", respondió Laurel.

2. Estime la proporción y el error estándar de la proporción para las órdenes de compra competitivas utilizando los datos de los archivos CH07A.xxx del CD que acompaña al libro. Elabore un intervalo de confianza del 95% para la proporción.

Bob observó escéptico los resultados. "¿Existe alguna manera de reducir esos límites del intervalo de confianza?", preguntó.

"Sin hacer un esfuerzo adicional de muestreo, estamos limitados a disminuir el nivel de confianza", explicó Laurel.

3. Calcule los límites del intervalo de confianza si Bob está dispuesto a contentarse con un nivel del 90%.

"La otra opción es emplear una muestra más grande", continuó ella. "Como el muestreo, en este caso, es relativamente poco costoso, ¿por qué no intentamos obtener un intervalo más pequeño, digamos más menos 0.03? Podemos utilizar nuestra proporción inicial como nuestra "valor fundamentado" con respecto a la proporción verdadera de la población y mantener nuestro nivel de confianza del 95%."

"¿Pero qué tanto más grande deberá ser la muestra?", preguntó Bob. "Te lo diré en un segundo", respondió Laurel al tiempo que sacaba su calculadora.

4. Con estas nuevas condiciones, ¿cuántas órdenes de adquisición más necesitan examinar?

"Buenas noticias", anunció Bob a Laurel varios días más tarde. "La nueva muestra más grande arrojó una proporción de 0.58. Eso significa que puedo decirle al jefe que estamos entre 0.55 y 0.61 con una certeza del 95%. Estoy planeando hacer una pequeña presentación juntos para el día de la reunión de la junta directiva." "Suena bien", dijo Laurel, "solamente ten cuidado en la forma en que utilizas los términos. Recuerda que hicimos unas cuantas triquiñuelas estadísticas en nuestros cálculos y no sería bueno que les causes una mala impresión".

5. Verifique los cálculos de Bob. ¿Qué piensa acerca de la preocupación de Laurel? ¿Cómo enfocaría la presentación si fuera Bob?

La presentación de Bob salió bien en la junta directiva del lunes siguiente. Hal hizo unas cuantas preguntas, pero en general se mostró complacido con los resultados. Luego pasó al siguiente punto a tratar.

"Como la mayoría de ustedes saben, hace aproximadamente un año introdujimos en nuestro inventario refacciones métricas. Con el flujo de equipo hidráulico portátil fabricado en el extranjero por compañías como Toyota, Nissan y Komatsu, el mercado de refacciones métricas parece estar maduro. Y hasta donde yo sé, fuimos los primeros en nuestro ramo en tener varias líneas completas. En cualquier caso, es hora de que veamos cómo estamos y de estimar las ventas potenciales para el año siguiente. Laurel, me temo que no te dejaremos descansar mucho, pero puedes darte cuenta que ¡definitivamente te necesitamos aquí!"

De regreso a su oficina, Laurel se puso a revisar lo que sabía de las líneas de refacciones métricas de HH Industries. Peggy estaba en proceso de pasarle un informe que le daría los detalles sobre las ventas del año anterior. Desafortunadamente, cuando se incorporaron las refacciones métricas, no se les asignó un código único de producto, lo cual hizo un tanto difícil aislar las ventas. Sin embargo, Laurel hizo lo que pudo.

6. Basándose en los datos de los archivos CH07B.xxx del CD que acompaña al libro, estime la media de la población y la desviación estándar de las ventas de refacciones métricas por semana.

7. Estime el error estándar de la media para esta muestra.

8. Construya un intervalo de confianza del 95% para las ventas semanales medias de refacciones métricas.

9. ¿Deberá HH Industries continuar ofreciendo refacciones métricas si Hal desea tener el 95% de confianza de que las ventas del año siguiente sean de al menos $300,000? Suponga que habrá 50 semanas hábiles durante el siguiente año.

l0. Stan argumentó que el uso de los 12 meses de datos sobre ventas de refacciones métricas daba una estimación demasiado baja, porque incluía los meses en que fueron introducidas. Está convencido que el uso de los datos correspondientes a los segundos seis meses mostrarán una predicción más precisa, ya que las ventas se habrían nivelado. Laurel está de acuerdo. Repita los cálculos anteriores sólo con los datos de las segundas 25 semanas.

Del libro de texto al mundo real

Fondo de Ingeniería en Berkeley*

Establecido en 1979, el Fondo de Ingeniería en Berkeley solicita contribuciones para apoyar al Colegio de Ingenieros de la Universidad de California, en Berkeley. Los administradores utilizan la información disponible acerca del número de donaciones, regalos y contribuciones en efectivo como entrada de un modelo matemático que predice las contribuciones al mes y al final del año. De acuerdo con la información obtenida ajustan los esfuerzos de obtención de fondos. El modelo utiliza una distribución binomial para la cantidad de donaciones y regalos, y una distribución de Poisson compuesta para la cantidad de dinero donada. Desde 1982, han registrado los datos de las cuentas de los donadores, periodicidad de las donaciones, tamaño de las donaciones, y la información equivalente de los regalos que hacen padres de familia, exalumnos, académicos y los amigos del Colegio.

Estimación de parámetros Los pronósticos están basadas en datos tomados de campañas anteriores. Como desde 1982 a 1984 se usó la misma correspondencia, las proporciones mensuales de las donaciones totales han sido estables de año en año. Para cada fecha de envío postal, los encargados de pronósticos determinan distribuciones para el número de donaciones de cada uno de los cuatro subgrupos, así como las estimaciones de la media y la varianza de las cantidades donadas.

Evaluación del modelo Los datos sobre los padres de familia, de 1982-1983 y 1983-1984 se utilizaron para probar la suposición de Poisson sobre la que se basa el modelo. Utilizando tanto las tablas de Poisson como una aproximación normal, se calcularon intervalos de confianza del 95% para el número de donaciones hechas por padres de familia. Las figuras MR7-1 y MR7-2 muestran estos intervalos para 1982-1983 y 1983-1984. Sólo en septiembre de ambos años las cuentas reales de los donadores cayeron fuera de los intervalos de confianza del 95%. Esto apoya la suposición de que se trata de una distribución de Poisson.

Resultados El modelo funcionó bien para pronosticar totales de fin de año, pero su desempeño fue un poco menor para los pronósticos mensuales. Las predicciones de las cuentas de donadores y de donaciones totales fueron más precisas para los padres, académicos y grupos de amigos que en el caso de los exalumnos. Los administradores pudieron entender mejor los efectos de los contactos personales y de los envíos por correo. Debido a que el modelo proporcionó una manera de predecir los efectos de los cambios en las técnicas de recaudación de fondos, los administradores se animaron a diseñar estrategias dirigidas a los grupos específicos.

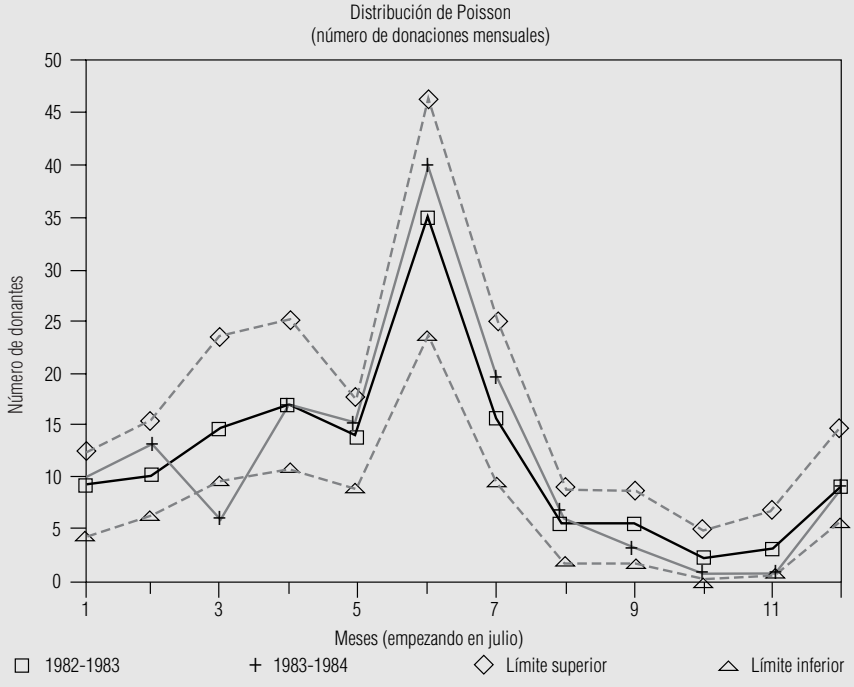


FIGURA MR7-1

Cuentas de las donaciones mensuales hechas por padres de familia durante 1982-1983

* *Fuente:* Mark Britto y Robert M. Oliver, "Forecasting Donors and Donations", *Journal of Forecasting* 5(1986): 39-55.

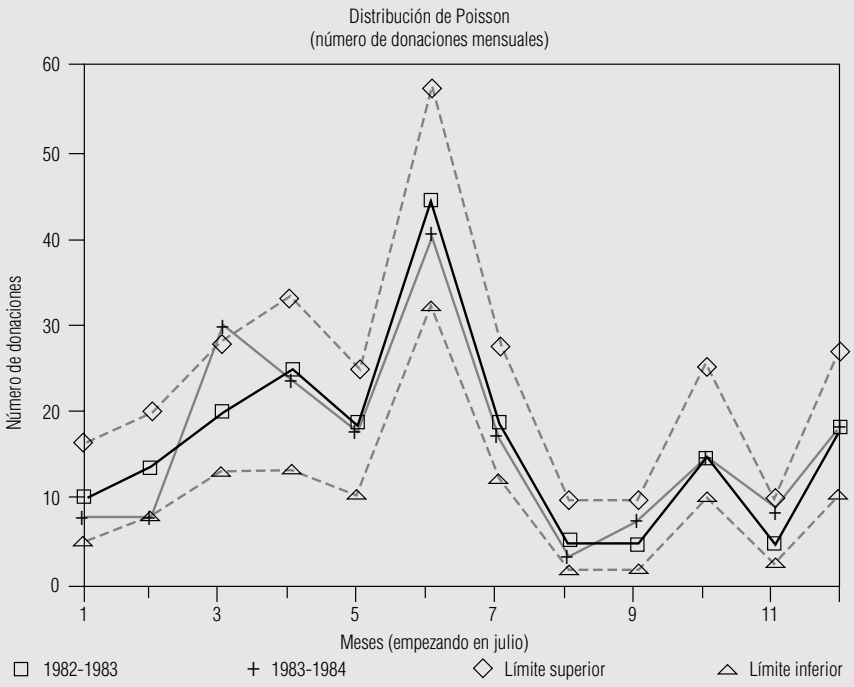


FIGURA MR7-2

Cuentas de las donaciones mensuales hechas por padres durante 1983-1984

Repaso del capítulo

● Términos introducidos en el capítulo 7

Distribución *t* de Student Familia de distribuciones de probabilidad que se distinguen por sus grados de libertad individuales; es parecida, en forma, a la distribución normal y se utiliza cuando se desconoce la desviación estándar de la población y el tamaño de la muestra es relativamente pequeño ($n \leq 30$).

Estimación Valor específico observado de un estimador.

Estimación de intervalo Un rango de valores utilizado para estimar un parámetro de población desconocido.

Estimación puntual Un solo número que se utiliza para estimar un parámetro de población desconocido.

Estimador Estadístico de muestra utilizada para estimar un parámetro de población.

Estimador consistente Estimador que produce valores que se acercan más al parámetro de la población conforme aumenta el tamaño de la muestra.

Estimador eficiente Estimador con un error estándar menor que algún otro estimador del parámetro de la población, esto es, cuanto más pequeño sea el error estándar de un estimador, más eficiente será.

Estimador no sesgado Estimador de un parámetro de población que, en promedio, toma valores mayores que el parámetro de la población con la misma frecuencia, y al mismo grado, con que tiende a tomar valores menores que el parámetro de la población.

Estimador suficiente Estimador que utiliza toda la información disponible en los datos correspondientes a un parámetro.

Grados de libertad Número de valores de una muestra que podemos especificar libremente, una vez que se sabe algo sobre dicha muestra.

Intervalo de confianza Un rango de valores que tiene designada una probabilidad de que incluya el valor verdadero del parámetro de la población.

Límites de confianza Límites inferior y superior de un intervalo de confianza.

Nivel de confianza Probabilidad que los estadísticos asocian a una estimación de intervalo de un parámetro y que indica qué tan seguros están de que la estimación de intervalo incluirá al parámetro de la población.

● Ecuaciones introducidas en el capítulo 7

■ **7-1** Estimación de la desviación estándar de la población

$$\hat{\sigma} = s = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}}$$

Esta fórmula indica que la desviación estándar de la muestra puede utilizarse para estimar la desviación estándar de la población.

■ **7-2**

$$\hat{\sigma}_{\bar{x}} = \frac{\hat{\sigma}}{\sqrt{n}} \times \sqrt{\frac{N - n}{N - 1}}$$

Esta fórmula nos permite derivar un error estándar *estimado* de la media de una población *finita* a partir de una *estimación* de la desviación estándar de la población. El símbolo ^ llamado gorro, indica que el valor es una estimación. La ecuación 7-6 es la fórmula correspondiente para una población infinita.

■ **7-3**

$$\mu_{\hat{p}} = p$$

Utilice esta fórmula para derivar la *media* de la distribución de muestreo de la *proporción* de éxitos. La parte derecha, p, es igual a $(n \times p)/n$, en donde el numerador es el número esperado de éxitos en n ensayos, y el denominador es el número de ensayos. En símbolos, la proporción de éxitos de *una muestra* se escribe como $\hat{p}$ y se lee *p gorro*.

■ **7-4**

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$$

Para obtener el *error estándar de la proporción,* obtenga la raíz cuadrada del producto de las probabilidades de éxito y de fracaso dividido entre el número de ensayos.

■ **7-5**

$$\hat{\sigma}_{\hat{p}} = \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

Ésta es la fórmula que se utiliza para derivar un error estándar estimado de la proporción, cuando se desconoce la proporción de la población y uno se ve forzado a utilizar $\bar{p}$ y $\bar{q}$, las proporciones de la muestra de éxitos y fracasos.

■ **7-6**

$$\hat{\sigma}_{\bar{x}} = \frac{\hat{\sigma}}{\sqrt{n}}$$

Esta fórmula nos permite derivar un error estándar *estimado* de la media de una población *infinita* a partir de una *estimación* de la desviación estándar de la población. Es bastante parecida a la ecuación 7-2, excepto porque carece del multiplicador de población finita.

● Ejercicios de repaso

■ **7-59** Para una muestra de 42 gasolineras en todo el estado, el precio promedio de un galón de gasolina sin plomo es $1.12 y la desviación estándar es $0.04 por galón. ¿Para qué intervalo puede tenerse el 99.74% de confianza de que incluirá la media estatal verdadera del precio por galón de gasolina sin plomo?

■ **7-60** ¿Cuáles son las ventajas de utilizar una estimación de intervalo en lugar de una estimación puntual?

■ **7-61** ¿Por qué es importante el error estándar de un estadístico cuando se utiliza como estimador? ¿Con qué característica de los estimadores se relaciona esto?

■ **7-62** Suzanne Jones, secretaria general del sistema universitario, necesita saber qué proporción de estudiantes tienen promedios de calificación menores que 2.0. ¿Cuántas calificaciones de estudiantes debe revisar con el fin de determinar la proporción que busca dentro de ± 0.01 con una confianza del 95%?

■ **7-63** Un intervalo de confianza del 95% para la media de la población está dado por (94, 126) y un intervalo de confianza del 75% está dado por (100.96, 119.04). ¿Cuáles son las ventajas y desventajas de cada una de estas estimaciones de intervalo?

■ **7-64** El límite de velocidad establecido en el Cross-Bronx Expressway es 55 mph. La congestión hace que la velocidad real sea mucho menor. Una muestra aleatoria de 57 vehículos dio un promedio de 23.2 mph y una desviación estándar de 0.3 mph.
a) Estime la desviación estándar de la población.
b) Estime el error estándar de la media para esta población.
c) ¿Cuáles son los límites superior e inferior del intervalo de confianza para la velocidad media dado un nivel de confianza deseado de 0.95?

■ **7-65** Dada una media de la muestra de 8, una desviación estándar de la población de 2.6 y una muestra de tamaño 32, encuentre el nivel de confianza asociado con cada uno de los siguientes intervalos:
a) (7.6136, 8.3864).
b) (6.85, 9.15).
c) (7.195, 8.805).

■ **7-66** Basándose en el conocimiento acerca de las cualidades deseables de los estimadores, ¿por qué razones debe considerarse a $\bar{x}$ como el "mejor" estimador de la media verdadera de la población?

■ **7-67** El presidente de la Offshore Oil ha estado preocupado acerca del número de peleas ocurridas en las instalaciones a su cargo y está considerando varios cursos de acción. En un esfuerzo por entender qué causa las peleas en alta mar, tomó una muestra aleatoria de 41 días en los que un equipo de trabajadores regresa a trabajar después de un permiso para ir a tierra firme. Para esta muestra, la proporción promedio de trabajadores que intervinieron en peleas cada día es 0.032, y la desviación estándar asociada es 0.0130.
a) Dé una estimación puntual de la proporción promedio de trabajadores que intervinieron en peleas en un día cualquiera en que la planta de trabajadores regresa de tierra firme.
b) Estime la desviación estándar de la población asociada con este índice de peleas.
c) Encuentre un intervalo de confianza del 90% para la proporción de trabajadores que regresan e intervienen en peleas.

■ **7-68** Dadas las siguientes expresiones para los límites de un intervalo de confianza, encuentre el nivel de confianza asociado con el intervalo:
a) $\bar{x} - 1.25\sigma_{\bar{x}}$ a $\bar{x} + 1.25\sigma_{\bar{x}}$.
b) $\bar{x} - 2.4\sigma_{\bar{x}}$ a $\bar{x} + 2.4\sigma_{\bar{x}}$.
c) $x - 1.68\sigma_{\bar{x}}$ a $\bar{x} + 1.68\sigma_{\bar{x}}$.

■ **7-69** La empresa Harris Polls, Inc., se dedica a investigar amas de casa. De encuestas anteriores, se sabe que la desviación estándar del número de horas por semana que un ama de casa dedica a ver televisión es de 1.1 horas. Harris Polls desea determinar el número promedio de horas por semana que un ama de casa en Estados Unidos dedica a ver televisión. La precisión es importante y, en consecuencia, Harris Polls quiere tener una certeza del 98% de que el número de muestra promedio de horas caerá dentro de ± 0.3 horas del promedio nacional. Conservadoramente, ¿qué tamaño de muestra deberá utilizar Harris Polls?

■ **7-70** John Bull acaba de adquirir un programa de computación que afirma escoger acciones que aumentarán su precio durante la semana siguiente con un índice de precisión del 85%. ¿En cuántas acciones deberá John probar el programa con el fin de estar el 98% seguro de que el porcentaje de acciones que realmente subirán de precio la semana próxima estará dentro de ± 0.05 de la proporción de la muestra?

■ **7-71** Gotchya es un centro de entretenimiento con instrumentos láser donde adultos y adolescentes rentan equipo y se enfrentan en un combate simulado. La instalación se usa a toda su capacidad los fines de semana. Los tres dueños quieren evaluar la efectividad de una nueva campaña de publicidad dirigida a aumentar su utilización entre semana. El número de clientes en 27 noches aleatorias entre semana está dado en la siguiente tabla. Encuentre un intervalo de confianza del 95% para el número medio de clientes en una noche entre semana.

61	57	53	60	64	57	54	58	63
59	50	60	60	57	58	62	63	60
61	54	50	54	61	51	53	62	57

7-72 Los contadores de Gotchya, el centro de entretenimiento del ejercicio 7-71, han informado a los dueños que necesitan tener al menos 55 clientes para salir *a mano* en una noche entre semana. Los socios están dispuestos a continuar operando entre semana si pueden tener una certeza del 95% o más de que saldrán a mano, al menos la mitad del tiempo. Use los datos del ejercicio 7-71 para encontrar un intervalo de confianza del 95% para la proporción de noches entre semana en que Gotchya saldrá a mano. ¿Deben continuar operando entre semana? Explique.

The Wall Street Journal proporciona información financiera diariamente respecto a más de 3,000 fondos de inversión mutua. La tabla MR7-1 da información de una muestra aleatoria de 35 de ellos y su desempeño al cierre del viernes 14 de mayo de 1993. Emplee esta información contestar los ejercicios del 7-73 al 7-76.

7-73 a) Estime el cambio promedio en el valor del activo neto (ΔVAN) del 14 de mayo de 1993 para todos los fondos listados en *The Wall Street Journal.*
b) Estime la desviación estándar del cambio en el valor del activo neto para todos los fondos del inciso a).
c) Encuentre un intervalo de confianza del 95% para el cambio promedio en el valor del activo neto. ¿Qué suposiciones necesita hacer acerca de la distribución del cambio individual en el valor del activo neto, con el fin de derivar el intervalo de confianza?

7-74 a) Estime la desviación estándar del cambio porcentual actualizado en valor (%ACT), de todos los fondos listados.
b) Suponiendo que la desviación estándar que estimó en el inciso a) es cercana a la desviación estándar real de la población, ¿qué tan grande deberá ser una muestra para estimar el cambio porcentual promedio actualizado en valor, dentro de 0.5% con el 99% de confianza?

7-75 Los fondos para los cuales el precio de oferta (PO) es el mismo que el valor del activo neto (VAN) se conocen como fondos "no cargados". Utilice la muestra de 35 fondos para estimar qué fracción de todos los fondos listados en *The Wall Street Journal* son fondos no cargados. Dé un intervalo de confianza del 98% para esta fracción.

7-76 Usted cree que los fondos no cargados no deberían agruparse con los demás. Suponiendo que los cambios porcentuales actualizados individuales en valor para los fondos no cargados tienen una distribución aproximadamente normal, encuentre un intervalo de confianza del 95% para su cambio porcentual promedio actualizado en valor. ¿Es necesario suponer la distribución normal? Explique su respuesta.

■ **7-77** Al evaluar la efectividad de un programa federal de rehabilitación, en una investigación de 52 de los 900 internos de una prisión se encontró que el 35% de éstos era reincidente.
a) Estime el error estándar de la proporción de reincidentes.
b) Construya un intervalo de confianza del 90% para la proporción de reincidentes entre los internos de esta prisión.

■ **7-78** Durante la cosecha de manzanas, se revisaron por separado 150 fanegas de la fruta en busca de manzanas en mal estado (debido, como usted sabe, a que una manzana mala puede echar a perder a todo el canasto) y se encontró que había un promedio de 3.2 manzanas malas por fanega. Se sabe que la desviación estándar de manzanas malas por fanega es de 0.2 para este tipo de manzana.
a) Calcule el error estándar de la media.
b) Establezca una estimación de intervalo alrededor de la media, utilizando una $\hat{\sigma}_{\bar{x}}$.

■ **7-79** De una muestra aleatoria de 60 autobuses, la oficina de transporte colectivo de la ciudad de Montreal ha calculado que el número medio de pasajeros por kilómetro es 4.1. De estudios anteriores se sabe que la desviación estándar de la población es 1.2 pasajeros por kilómetro.
a) Encuentre el error estándar de la media. (Suponga que la flotilla de autobuses es muy grande.)
b) Construya un intervalo de confianza del 95% para el número medio de pasajeros por kilómetro para la población.

■ **7-80** Recientemente, el Servicio de Impuestos de Estados Unidos tomó una muestra de 200 devoluciones de impuestos y encontró que el reembolso promedio de impuestos de la muestra llegaba a $425.39, con una desviación estándar de la muestra de $107.10.
a) Estime el reembolso medio de impuestos y la desviación estándar de la población.
b) Utilizando las estimaciones hechas en el inciso anterior, construya un intervalo con el 95% de certeza de que la media de la población estará en él.

■ **7-81** Physicians Care Group opera varias clínicas que atienden sin cita. Los expedientes de los pacientes indican la hora en que llega a la clínica y la hora en que un médico atiende a ese paciente. El administrador Val Likmer acaba de recibir una desagradable llamada telefónica de un paciente que se quejó de una espera excesiva en la clínica de Rockridge. Val saca 49 expedientes al azar de la semana pasada y calcula un tiempo de espera promedio de 15.2 minutos. Un estudio anterior de gran escala del tiempo de espera en varias clínicas obtuvo una desviación estándar de 2.5 minutos. Elabore un intervalo de confianza para el tiempo de espera promedio con nivel de confianza del
a) 90%.
b) 99%.

■ **7-82** Bill Wenslaff, un ingeniero de una planta purificadora de agua, mide diariamente el contenido de cloro en 200 muestras diferentes. En un periodo de varios años, ha establecido que la desviación estándar de la po-

Tabla MR7-1

Datos financieros para una muestra de 35 fondos mutuos

Nombre del fondo	VAN	PO	ΔVAN	%ACT
AHA Balanced	12.54	12.54	−0.01	3.9
Ambassador Index Stock	11.36	11.36	0.01	1.9
American Capital Global Equity (A)	10.44	11.08	0.01	8.2
American Capital Municipal Bond	10.33	10.85	−0.01	5.1
Atlas Growth & Income	13.69	14.04	−0.05	2.2
Babson Enterprise	16.13	16.13	0.08	6.0
Blanchard Flexible Income	5.11	5.11	0.00	5.9
Colonial Growth	14.08	14.94	−0.05	0.1
Columbia Common Stock	14.54	14.54	−0.02	3.8
Evergreen Total Return	19.96	19.96	−0.07	5.9
Fidelity Equity-Income	31.24	31.88	−0.14	8.6
Fidelity Spartan Municipal Income	11.02	11.02	0.00	5.9
First Union Value (B)	17.30	18.02	−0.04	1.8
Flag Investors Value	10.89	11.40	−0.05	2.9
Fortis Capital	17.48	18.35	0.03	−5.3
GT Global Europe	9.11	9.56	0.03	7.1
Helmsman Equity Index	11.68	11.68	0.02	1.8
Homestead Value	13.48	13.48	−0.01	7.9
IAI Emerging Growth	13.64	13.64	0.09	−2.8
John Hancock Tax Exempt	11.32	11.85	0.00	5.1
Kemper Blue Chip	13.30	14.11	0.02	−0.2
Keystone International	6.50	6.50	0.01	8.0
Marshall Stock	9.90	9.90	0.03	−1.9
MAS Equity	54.37	54.37	−0.11	−1.9
MFS Research	12.86	13.64	0.01	4.6
MIM Bond Income	9.24	9.24	0.02	−0.5
PFAMCo MidCap Growth	12.51	12.51	−0.03	2.8
Pilgrim GNMA	14.02	14.45	−0.01	3.2
PIMCO Short Term	10.03	10.03	0.01	1.8
Prudential Municipal Maryland	11.35	11.35	0.00	4.8
Putnam Global Growth	8.18	8.68	−0.01	10.1
Rightime Blue Chip	31.07	32.62	0.02	1.2
Schwab 1000	12.11	12.11	−0.01	1.3
Shearson Appreciation (A)	10.72	11.28	−0.03	0.6
Weiss Peck Greer Tudor	24.90	24.90	0.19	0.2

VAN Valor del activo neto, precio (en dólares) al cual un inversionista puede redimir acciones del fondo.

PO Precio de oferta, precio (en dólares) que paga un inversionista para adquirir acciones del fondo.

ΔVAN Cambio en el VAN respecto al día anterior.

%ACT Cambio porcentual actualizado en el valor de una inversión en el fondo, suponiendo que todos los dividendos se reinvierten.

Fuente: The Wall Street Journal *(17 de mayo de 1993), págs. C16-C19.*

blación es de 1.4 miligramos de cloro por litro. Las últimas muestras arrojaron un promedio de 4.6 miligramos de cloro por litro.

a) Encuentre el error estándar de la media.
b) Establezca el intervalo alrededor de 5.2, la media de la población, que incluirá a la media de la muestra con una probabilidad del 68.3%.

■ **7-83** Ellen Harris, una ingeniera industrial, estuvo acumulando tiempos normales para varias tareas sobre un proceso de ensamble de trabajo intensivo. Este proceso incluía 300 estaciones de trabajo diferentes, cada una efectuando las mismas actividades de ensamble. Muestreó siete estaciones y obtuvo los siguientes tiempos de ensamble, en minutos, para cada estación: 1.9, 2.5, 2.9, 1.3, 2.6, 2.8 y 3.0.

a) Calcule el tiempo medio de ensamble y la desviación estándar correspondiente para la muestra.
b) Estime la desviación estándar de la población.
c) Dé un intervalo de confianza del 98% para el tiempo medio de ensamble.

■ **7-84** Larry Culler, inspector federal de granos en un puerto marítimo, encontró que había partes echadas a perder en 40 de 120 lotes de avena, elegidos aleatoriamente, embarcados en el puerto. Construya un intervalo de confianza del 95% para la proporción real de lotes con partes echadas a perder en embarques hechos desde ese puerto.

■ **7-85** La compañía de confección de ropa High Fashion Marketing está considerando la recolocación en el mercado de corbatas de lana de cachemira. Con el fin de evitar un fracaso, la High Fashion entrevistó a 90 jóvenes ejecutivos (su principal mercado) y encontró que de los 90 entrevistados, 79 creían que las corbatas de cachemira estaban de moda y les interesaba comprarse una. Use un nivel de confianza del 98% para dar un intervalo de confianza para la proporción de todos los jóvenes ejecutivos que piensan que las corbatas de cachemira están de moda.

■ **7-86** El Departamento de Transporte ha ordenado que la velocidad promedio de los automóviles en la carretera interestatal no debe sobrepasar las 67 millas por hora, para que los departamentos de carreteras del estado puedan retener su presupuesto federal. Agentes de la policía de caminos de Carolina del Norte, en automóviles sin insignias, tomaron una muestra de 186 coches y encontraron que la velocidad promedio era 66.3 millas por, con una desviación estándar de 0.6 millas por hora.

a) Encuentre el error estándar de la media.
b) ¿Cuál es el intervalo alrededor de la media de la muestra que contendría a la media de la población el 95.5% de las veces?
c) ¿Puede el departamento de transporte de Carolina del Norte informar con veracidad que la velocidad promedio real de sus carreteras es 67 millas por hora o menos con el 95.5% de confianza?

■ **7-87** Mark Semmes, dueño del restaurante Aurora, está considerando la compra de nuevo mobiliario. Como ayuda para decidir sobre la cantidad que puede invertir en mesas y sillas, desea determinar el ingreso por cliente. Tomó una muestra aleatoria de nueve clientes, cuyo consumo promedio fue $18.30 con una desviación estándar de $3.60. Elabore un intervalo de confianza del 95% para la cantidad promedio por cliente en la nota de consumo.

■ **7-88** John Deer, un horticultor de la Universidad Estatal de Northern Carrboro, sabe que cierta especie de maíz siempre produce entre 80 y 140 fanegas por hectárea. Para un nivel de confianza del 90%, ¿cuántas muestras de una hectárea debe tomar con el fin de estimar la producción promedio por hectárea dentro de ± 5 fanegas por hectárea?

capítulo 8

PRUEBA DE HIPÓTESIS: PRUEBA DE UNA SOLA MUESTRA

Objetivos

- Aprender cómo usar las muestras para decidir si una población posee una característica dada
- Determinar qué tan improbable es que una muestra observada provenga de una población hipotética
- Comprender los dos tipos de errores posibles que se producen al probar las hipótesis
- Aprender cuándo usar pruebas de una cola y cuándo pruebas de dos colas
- Aprender el proceso de cinco pasos para probar hipótesis
- Aprender cómo y cuándo usar las distribuciones *t* y normal para probar hipótesis sobre medias y proporciones de población

Contenido del capítulo

8.1 Introducción 320
8.2 Conceptos básicos en el procedimiento de prueba de hipótesis 321
8.3 Prueba de hipótesis 324
8.4 Prueba de hipótesis de medias cuando se conoce σ la desviación estándar de la población 331
8.5 Medición de la potencia de una prueba de hipótesis 338
8.6 Prueba de hipótesis para proporciones: muestras grandes 341
8.7 Prueba de hipótesis de medias cuando no se conoce la desviación estándar de la población 347
- Estadística en el trabajo 351
- Ejercicio de base de datos computacional 351
- Del libro de texto al mundo real 352
- Términos introducidos en el capítulo 8 353
- Ejercicios de repaso 353

El contrato para techar un nuevo complejo deportivo de San Francisco fue otorgado a Parkhill Associates, una compañía constructora grande. Las especificaciones de construcción indican un techo móvil de aproximadamente 10,000 láminas de aluminio con un grosor de 0.04 pulgadas (pulg). Las láminas de aluminio no pueden ser considerablemente más gruesas que 0.04 pulg porque la estructura podría no soportar el peso adicional. Y tampoco pueden ser más delgadas porque la fuerza estructural del techo sería inadecuada. Debido a estas restricciones en el grosor, Parkhill revisa cuidadosamente las láminas de aluminio de su proveedor. Es claro que Parkhill no desea medir cada lámina, así que toma una muestra aleatoria de 100. Las láminas de la muestra tienen un grosor medio de 0.0408 pulg. Por experiencia con este proveedor, Parkhill cree que esas láminas provienen de una población que tiene una desviación estándar de 0.004 pulg de grosor. Con base en estos datos, Parkhill debe decidir si las 10,000 láminas cumplen con las especificaciones. En el capitulo 7, utilizamos estadísticas muestrales para estimar parámetros de población. Ahora, para resolver problemas como el de Parkhill, aprenderemos cómo utilizar las características de las muestras para probar una suposición acerca de la población de la que se tomó dicha muestra. Nuestra prueba para Parkhill, más adelante en este capítulo, llevará a que la constructora acepte el cargamento de láminas o que lo rechace por no cumplir con las especificaciones arquitectónicas. ■

8.1 Introducción

Función de la prueba de hipótesis

La prueba de hipótesis comienza con una suposición, llamada *hipótesis*, que hacemos acerca de un parámetro de población. Después recolectamos datos de muestra, producimos estadísticas muestrales y usamos esta información para decidir qué tan probable es que nuestro parámetro de población hipotético sea correcto. Digamos que suponemos un cierto valor para una media de población. Para probar la validez de esa suposición recolectamos datos de muestra y determinamos la diferencia entre el valor hipotético y el valor real de la media de la muestra. Después juzgamos si la diferencia obtenida es significativa o no. Mientras más pequeña sea la diferencia, mayor será la probabilidad de que nuestro valor hipotético para la media sea correcto. Mientras mayor sea la diferencia, más pequeña será la probabilidad.

Desafortunadamente, la diferencia entre el parámetro de población hipotético y la estadística real rara vez es tan grande que rechacemos automáticamente nuestra hipótesis o tan pequeña que la aceptamos con la misma rapidez. Así, en las pruebas de hipótesis como en la mayoría de las decisiones importantes de la vida real, las soluciones claras o bien definidas son la excepción, no la regla.

Cuándo aceptar o rechazar la hipótesis

Supongamos que una administradora de un gran centro comercial nos dice que la eficiencia de trabajo promedio de sus empleados es de al menos 90%. ¿Cómo podemos probar la validez de su hipótesis? Utilizando los métodos de muestreo aprendidos en el capítulo 6, podríamos calcular la eficiencia de una *muestra* de sus empleados. Si hiciéramos esto y el estadístico de la muestra resultara ser 95%, aceptaríamos sin demora la afirmación de la administradora. Sin embargo, si el estadístico de la muestra resultara ser el 46%, rechazaríamos su afirmación por falsa. Podemos interpretar estos dos resultados, 95 y 46%, si utilizamos el sentido común.

El problema básico es manejar la incertidumbre

Ahora supongamos que nuestro estadístico revela una eficiencia del 88%. Este valor es relativamente cercano al 90%. Pero, ¿es suficientemente cercano para que aceptemos como correcta la hipótesis de la administradora? Ya sea que aceptemos o rechacemos su hipótesis, no podemos estar ab-

solutamente seguros de que nuestra decisión es correcta; por consiguiente, tendremos que aprender cómo manejar la incertidumbre en nuestra toma de decisiones. **No podemos aceptar o rechazar una hipótesis sobre un parámetro de población sólo por intuición. Más bien, necesitamos aprender cómo decidir objetivamente si aceptamos o rechazamos una corazonada, con base en la información de la muestra.**

Se sacan grandes conclusiones

Proyección excesiva

Los estudiantes universitarios a menudo buscan anuncios que ofrecen ayuda para el aprendizaje. Una de las más populares de estas ayudas es una combinación de descripciones breves, una guía de estudio y un juego de preguntas para varios cursos. Los anuncios de estos artículos, con frecuencia, afirman que se pueden obtener mejores calificaciones en los exámenes con menos tiempo de estudio. Supongamos que una guía de estudio para un curso de estadística básica está disponible a través de una organización que produce tales guías para 50 cursos diferentes. Si esta guía de estudio para estadística básica ha sido probada (y supongamos que adecuadamente), la compañía podría anunciar: "ha sido probado estadísticamente que nuestras guías de estudio aumentan las calificaciones y disminuyen el tiempo de estudio". Desde luego que esta afirmación es bastante cierta, pero sólo en lo que se refiere a la experiencia de estadística básica. Puede ser que no haya evidencia alguna con significado estadístico que establezca el mismo tipo de resultados para las otras 49 guías.

Condiciones de prueba diferentes

Otro producto puede anunciar que elimina la hierba silvestre que crece en el césped y puede afirmar que el producto ha sido "minuciosamente probado" en céspedes reales. Aun si suponemos que se utilizaron efectivamente los procedimientos estadísticos adecuados durante las pruebas, tales afirmaciones siguen siendo "grandes conclusiones". Supongamos que el terreno de prueba estaba en Florida, y el césped con problemas se encuentra en Utah. Las diferencias de lluvia, fertilidad del suelo, contaminantes del aire, temperatura, horas de inactividad y condiciones de germinación pueden variar ampliamente entre estos dos lugares. Afirmar ciertos resultados para una prueba estadísticamente válida para un conjunto de condiciones de prueba completamente diferente es algo que no posee valor alguno. Una prueba de este tipo no puede medir la efectividad en una amplia variedad de condiciones ambientales.

Ejercicios 8.1

- **8-1** ¿Por qué es necesario que manejemos la incertidumbre en nuestra toma de decisiones, incluso cuando usamos técnicas estadísticas?
- **8-2** Teóricamente, ¿cómo probaría la hipótesis de que una moneda no está alterada? ¿O de que un dado no está cargado?
- **8-3** ¿Es posible que se acepte una hipótesis falsa? ¿Cómo explicaría esto?
- **8-4** Describa el proceso de prueba de hipótesis.
- **8-5** ¿Cómo explicaría que se tuviera una gran diferencia entre un parámetro de población hipotético y un estadístico si, en efecto, la hipótesis es cierta?

8.2 Conceptos básicos en el procedimiento de prueba de hipótesis

Problema del complejo deportivo

Antes de introducir los términos y procedimientos estadísticos formales, analizaremos completamente nuestro problema del complejo deportivo que abre el capítulo. Recordemos que es necesario que las láminas de aluminio del techo tengan un grosor de 0.04 pulg y que serán insatisfactorias si resultan demasiado gruesas *o* demasiado delgadas. El contratista toma una muestra de 100 láminas y determina que el grosor medio de la muestra es 0.0408 pulg. Con base en su experiencia, sabe que la desviación estándar de la población es 0.004 pulg. ¿Esta evidencia de muestra indica que el lote de 10,000 láminas de aluminio es adecuado para construir el techo del nuevo complejo deportivo?

Formulación de la hipótesis

Si suponemos que el grosor medio real es 0.04 pulg y sabemos que la desviación estándar de la población es 0.004 pulg, ¿qué tan probable es que obtengamos una media de la muestra de 0.0408 o más a partir de la población? En otras palabras, **si la media real es 0.04 pulg y la desviación estándar es 0.004 pulg, ¿cuáles son las posibilidades de obtener una media de la muestra que difiera de 0.04 pulg en 0.0008 pulg (= 0.0408 − 0.04) o más?**

Estas preguntas muestran que **para determinar si la media de la población es realmente 0.04 pulg debemos calcular la probabilidad de que una muestra aleatoria con una media de 0.0408 pulg sea seleccionada de una población con una μ de 0.04 pulg y una σ de 0.004 pulg. Esta probabilidad indicará si es *razonable* observar una muestra como ésta si la media de población es realmente 0.04 pulg.** Si esta probabilidad es demasiado baja, debemos concluir que la afirmación de la compañía de aluminio es falsa y que el grosor medio de las láminas de aluminio no es 0.04 pulg.

Respondamos a la pregunta ilustrada en la figura 8-1: si la media de población hipotética es 0.04 pulg y la desviación estándar de la población es 0.004 pulg, ¿cuáles son las posibilidades de obtener una media de la muestra (0.0408 pulg) que difiera de 0.04 pulg en 0.0008 pulg? Primero, calculemos el error estándar de la media a partir de la desviación estándar de la población:

Cálculo del error estándar de la media

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \qquad [6\text{-}1]$$

$$= \frac{0.004 \text{ pulg}}{\sqrt{100}}$$

$$= \frac{0.004 \text{ pulg}}{10}$$

$$= 0.0004 \text{ pulg}$$

A continuación usamos la ecuación 6-2 para descubrir que la media de nuestra muestra (0.0408 pulg) cae dos errores estándar a la derecha de la media hipotética de la población:

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} \qquad [6\text{-}2]$$

$$= \frac{0.0408 - 0.04}{0.0004}$$

$$= 2 \leftarrow \text{Errores estándar de la media}$$

Interpretación de la probabilidad asociada con esta diferencia

Usando la tabla 1 del apéndice, sabemos que 4.5% es la probabilidad total de que nuestra media muestral difiera de la media de población en dos o más errores estándar, esto es, la probabilidad de que la media de la muestra sea 0.0408 pulg o mayor, o 0.0392 pulg o menor es sólo 4.5% ($P(z \geq 2$ o $z \leq -2) = 2(0.5 - 0.4772) = 0.0456$ o alrededor de 4.5%). **Con esta posibilidad tan baja, Parkhill podría concluir que no es probable que una población con una media real de 0.04 pulg produzca una muestra como ésta.** El supervisor del proyecto rechazaría la afirmación de la compañía de aluminio respecto al grosor medio de las láminas.

El papel del tomador de decisiones en la formulación de hipótesis

En este caso, la diferencia entre la media de la muestra y la media hipotética de la población es demasiado grande, y la probabilidad de que la población produzca una muestra aleatoria semejante es demasiado baja. Por qué esta probabilidad de 4.5% es demasiado baja, o errónea, es un juicio que deben hacer los tomadores de decisiones. Ciertas situaciones demandan que los tomadores de decisiones estén muy seguros respecto a las características de los elementos que se están probando, y en ese caso incluso 2% es demasiado alto para ser atribuible al azar. Otros procesos permiten una libertad o variación más amplia, y el tomador de decisiones puede aceptar una hipótesis con una probabilidad del 4.5% de variación aleatoria. En cada situación, es necesario determinar los costos resultantes de tomar una decisión incorrecta y el nivel preciso de riesgo que estamos dispuestos a correr.

Riesgo de rechazo

En nuestro ejemplo, rechazamos el argumento de la compañía de techos de aluminio que dice que la media de la población es 0.04 pulg. Pero, supongamos por un momento que la media de la pobla-

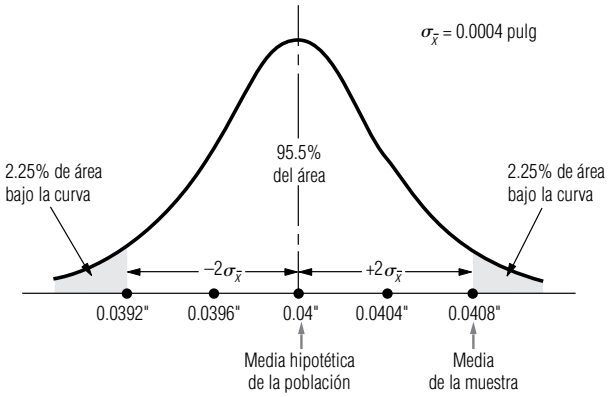


FIGURA 8-1

Probabilidad de que $\bar{x}$ difiera de la μ hipotética en 2 errores estándar

ción fuera *realmente* 0.04 pulg. Si nos atuviéramos a nuestra regla de rechazo de dos o más errores estándar (la probabilidad del 4.5% o menos en las colas de la figura 8-1), rechazaríamos un lote de láminas de aluminio en perfectas condiciones 4.5% de las veces. En consecuencia, **nuestro estándar mínimo para una probabilidad aceptable, 4.5%, *también* es el *riesgo* que corremos de *rechazar* una *hipótesis que es cierta.* En ésta o cualquier otra toma de decisiones, no puede haber transacciones libres de riesgo.**

SUGERENCIAS Y SUPOSICIONES

Aunque realizar *pruebas de hipótesis* se oye como un término estadístico formal sin relación con la toma de decisiones de negocios, de hecho los administradores proponen y prueban hipótesis todo el tiempo. "Si bajamos $1,500 el precio de este modelo de auto, venderemos 50,000 de ellos este año" es una hipótesis. Para probarla, será necesario esperar al final del año y contar las ventas. Las hipótesis administrativas se basan en la intuición; el mercado decide si esa intuición era correcta. Sugerencia: una prueba de hipótesis se refiere a hacer inferencias acerca de la población a partir de una muestra muy pequeña. Lo más importante de las pruebas de hipótesis es que nos permiten preguntarnos (y después decidir) si es probable que una población como la que *creemos* que es ésta producirá una muestra como la que buscamos.

Ejercicios 8.2

Ejercicios de autoevaluación

EA 8-1 ¿Cuántos errores estándar alrededor del valor hipotético deben usarse para estar 99.44% seguros de que se acepta una hipótesis cuando es cierta?

EA 8-2 Un fabricante de automóviles asegura que un modelo específico rinde 28 millas por galón. La Agencia de Protección Ambiental usó una muestra de 49 autos de este modelo y encontró un media de la muestra de 26.8 millas por galón. De estudios anteriores se sabe que la desviación estándar de la población es 5 millas por galón. ¿Es razonable esperar (dentro de 2 errores estándar) que se puede seleccionar esa muestra si en realidad la media verdadera de la población es 28 millas por galón?

Conceptos básicos

8-6 ¿Qué queremos decir cuando rechazamos una hipótesis con base en una muestra?

8-7 Explique por qué no se usa un solo nivel estándar de probabilidad para rechazar o aceptar en la prueba de hipótesis.

8-8 Si rechazamos un valor hipotético porque difiere de un estadístico de la muestra en más de 1.75 errores estándar, ¿cuál es la probabilidad de que hayamos rechazado una hipótesis que de hecho es cierta?

8-9 ¿Cuántos errores estándar alrededor del valor hipotético debemos usar para estar 98% seguros de que aceptamos la hipótesis cuando es cierta?

Aplicaciones

8-10 El magnate deportivo y de los medios Ned Sterner está interesado en comprar el equipo de los Stalwarts de Atlanta, si puede estar razonablemente seguro de que su manejo no será demasiado costoso. Piensa que la asistencia promedio aproximada tendría que ser de 28,500 aficionados por juego para que la compra le resultara atractiva. Ned elige al azar 64 juegos en casa entre los realizados durante los cuatro años anteriores y encuentra, de las cifras reportadas en *Sporting Reviews*, que la asistencia promedio a estos juegos fue de 26,100 aficionados. Un estudio que encargó anteriormente al comprar un equipo mostró que la desviación estándar de la población de asistentes a eventos similares había sido bastante estable durante los 10 años anteriores, en alrededor de 6,000 aficionados. Utilizando dos errores estándar como el criterio de decisión, ¿debería Ned comprar el equipo de los Stalwarts? ¿Puede pensar en alguna(s) razón(es) por las que su conclusión tal vez no sea válida?

8-11 La revista *Computing World* ha declarado que el propietario promedio de una computadora personal pasa 23.9 horas a la semana usando su máquina, con una desviación estándar de 12.6 horas a la semana. Un muestreo aleatorio de 81 suscriptores reveló una media muestral de 27.2 horas por semana. Con base en esta muestra, ¿es razonable concluir (usando dos errores estándar como criterio de decisión) que los suscriptores de *Computing World* son distintos de los propietarios promedio de computadoras personales?

8-12 Una tienda de abarrotes ha empacado naranjas en bolsas especiales y asegura que una bolsa rinde 2.5 litros de jugo. Después de seleccionar al azar 42 bolsas, el empacador encontró que la producción promedio de jugo por bolsa era 2.2 litros. Datos históricos establecen que la desviación estándar de la población es de 0.2 litros. Usando esta muestra y un criterio de decisión de 2.5 errores estándar, ¿se puede concluir que la afirmación de la tienda es correcta?

Soluciones a los ejercicios de autoevaluación

EA 8-1 Para dejar una probabilidad de $1 - 0.9944 = 0.0056$ en las colas, el valor absoluto de z debe ser mayor o igual que 2.77, de manera que el intervalo debe ser ± 2.77 errores estándar alrededor del valor hipotético.

EA 8-2 $\sigma = 5 \quad n = 49 \quad \bar{x} = 26.8 \quad \mu = 28$

$$\mu \pm 2\sigma_{\bar{x}} = \mu \pm 2\sigma/\sqrt{n} = 28 \pm 2(5)/\sqrt{49} = 28 \pm 1.429 = (26.571, 29.429)$$

Dado que $\bar{x} = 26.8 > 26.57$, es razonable obtener esos resultados de la muestra si μ es realmente 28 mpg.

8.3 Prueba de hipótesis

Declaración formal de la *hipótesis nula*

En una prueba de hipótesis, debemos establecer el valor supuesto o hipotético del parámetro de población *antes* de comenzar a tomar la muestra. La suposición que deseamos probar se conoce como *hipótesis nula* y se simboliza H_0, o "H sub-cero".

Supongamos que deseamos probar la hipótesis de que la media de la población es igual a 500. En símbolos se escribe como sigue y se lee "la hipótesis nula es que la media de población es igual a 500":

$$H_0: \mu = 500$$

¿Por qué se llama "hipótesis nula"?

El término *hipótesis nula* surge de las primeras aplicaciones agrícolas y médicas de la estadística. Con el fin de probar la efectividad de un nuevo fertilizante o de una nueva medicina, la hipótesis que se probaba era que *no hubo efecto*, es decir, no hubo diferencia entre las muestras tratadas y las no tratadas.

Si en un problema usáramos un valor hipotético de una media de población, en símbolos sería:

$$\mu_{H_0}$$

Esto se lee, "valor hipotético de la media de la población".

Si los resultados de nuestra muestra no respaldan la hipótesis nula, debemos concluir que se cumple alguna otra cosa. **Siempre que rechazamos la hipótesis, la conclusión que sí aceptamos se llama *hipótesis alternativa* cuyo símbolo es H_1 ("H sub-uno").** Para la hipótesis nula:

$$H_0: \mu = 200 \text{ (Se lee: "la hipótesis nula es que la media de población es igual a 200")}$$

consideraremos tres hipótesis alternativas posibles:

Declaración formal de la *hipótesis alternativa*

- H_1: $\mu \neq 200$ ← "La hipótesis alternativa es que la media de población *no es igual* a 200"
- H_1: $\mu > 200$ ← "La hipótesis alternativa es que la media de población *es mayor que* 200"
- H_1: $\mu < 200$ ← "La hipótesis alternativa es que la media de población *es menor que* 200"

Interpretación del nivel de significancia

Objetivo de la prueba de hipótesis

El propósito de la prueba de hipótesis no es cuestionar el valor calculado del estadístico de la muestra, sino hacer un juicio respecto a la *diferencia* entre ese estadístico y un parámetro hipotético de la población. El siguiente paso después de establecer las hipótesis nula y alternativa, entonces, consiste en decidir qué criterio utilizar para confirmar si se acepta o se rechaza la hipótesis nula.

En nuestro ejemplo del complejo deportivo, decidimos que una diferencia observada entre la media de la muestra $\bar{x}$ y la media hipotética de la población μ_{H_0} tenía sólo 4.5%, o 0.045, de probabilidad de ocurrir. Por tanto, rechazamos la hipótesis nula de que la media de la población fuera de 0.04 pulg (H_0: $\mu = 0.04$ pulg). En términos estadísticos, el valor 0.045 se conoce como *nivel de significancia*.

Función del *nivel de significancia*

¿Qué pasa si probamos una hipótesis con 5% de nivel de significancia? Esto quiere decir que rechazaremos la hipótesis nula si la diferencia entre el estadístico y el parámetro hipotético de la población es tan grande que ésta u otra diferencia mayor ocurrirá, en promedio, sólo cinco o menos veces en cada 100 muestras, cuando el parámetro hipotético de la población es correcto. **Si suponemos que la hipótesis es correcta, entonces el nivel de significancia indicará el porcentaje de medias muestrales que está fuera de ciertos límites.** (Recuerde que en estimación, el nivel de confianza indicaba el porcentaje de medias muestrales que caían *dentro* de los límites de confianza definidos.)

Área en la que no existe una diferencia significativa

La figura 8-2 ilustra cómo interpretar un nivel de significancia del 5%. Observe que el 2.5% del área bajo la curva está localizado en cada cola. De la tabla 1 del apéndice, podemos determinar que el 95% de toda el área bajo la curva está incluido en un intervalo que se extiende $1.96\sigma_{\bar{x}}$ a cada lado de la media hipotética. Así pues, en el 95% del área no hay diferencia significativa entre el valor observado del estadístico de la muestra y el valor hipotético del parámetro de población. En el 5% restante (las regiones sombreadas de la figura 8-2) sí existe una diferencia significativa.

También llamada área de aceptación de la hipótesis nula

La figura 8-3 examina este mismo ejemplo de otra manera. Aquí, 0.95 del área bajo la curva cae en la región donde aceptaríamos la hipótesis nula. Las dos partes sombreadas bajo la curva, que representan un total del 5% del área, son las regiones donde rechazaríamos la hipótesis nula.

Las hipótesis son *aceptadas,* no *demostradas*

Resulta conveniente hacer un advertencia en este punto. Aunque el estadístico de la figura 8-3 cae en la región no sombreada (la región que comprende 95% del área bajo la curva), **esto *no prueba* que nuestra hipótesis nula (H_0) sea cierta; simplemente no nos proporciona evidencia estadís-**

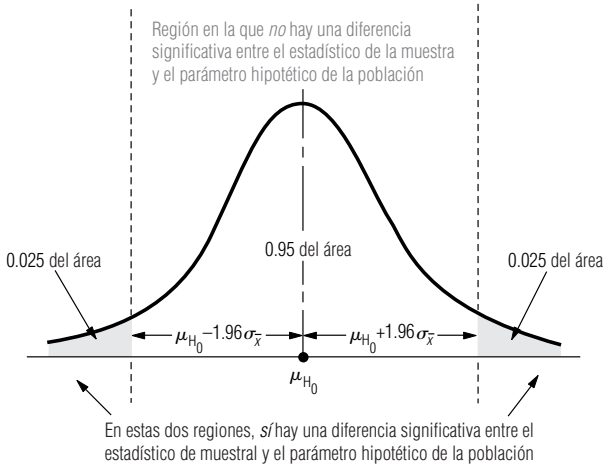


FIGURA 8-2

Regiones de diferencia significativa y de diferencia no significativa para un nivel de significancia del 5%

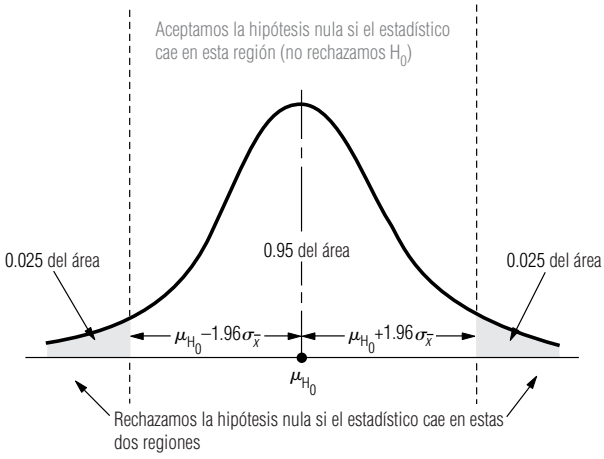


FIGURA 8-3

Un nivel de significancia del 5%, con regiones designadas de aceptación y rechazo

tica para rechazarla. ¿Por qué? Porque la única forma en que puede aceptarse la hipótesis con seguridad es que conozcamos el parámetro de población y, desafortunadamente, esto no es posible. Por consiguiente, siempre que afirmemos que aceptamos la hipótesis nula, en realidad lo que queremos decir es que no hay suficiente evidencia estadística para rechazarla. **El empleo del término *aceptar*, en vez de *no rechazar*, se ha convertido en el estándar. Significa simplemente que cuando los datos de la muestra no hacen que rechacemos una hipótesis nula, nos comportamos como si esa hipótesis fuera cierta.**

Selección de un nivel de significancia

Trueque al escoger un nivel de significancia

No existe un nivel de significancia único estándar o universal para probar hipótesis. En algunos casos, se utiliza un nivel de significancia del 5%. Ciertos resultados de investigaciones publicados a menudo prueban hipótesis para un nivel de significancia del 1%. Es posible probar una hipótesis a *cualquier* nivel de significancia. Pero recordemos que nuestra elección del estándar mínimo para una probabilidad aceptable, o el nivel de significancia, es también el riesgo que corremos al rechazar una hipótesis nula cuando es cierta. **Cuanto más alto sea el nivel de significancia que utilizamos para probar una hipótesis, mayor será la probabilidad de rechazar la hipótesis nula cuando es cierta.**

Al examinar este concepto, nos referiremos a la figura 8-4. Ésta ilustra una prueba de hipótesis para tres niveles de significancia: 0.01, 0.10 y 0.50. También hemos indicado la ubicación de la misma media de muestra $\bar{x}$ en cada distribución. En las gráficas (a) y (b), aceptaríamos la hipótesis nula de que la media de población es igual al valor hipotetizado. Pero observe que en la gráfica (c), rechazaríamos esta misma hipótesis nula. ¿Por qué? Nuestro nivel de significancia de 0.50 en esta parte es tan alto que rara vez aceptaríamos la hipótesis nula cuando *no* es cierta, pero, al mismo tiempo, con frecuencia la rechazaríamos aunque *lo sea*.

Errores de tipo I y tipo II

Definición de los errores tipo I y tipo II

Los especialistas en estadística usan definiciones y símbolos específicos para el concepto ilustrado en la figura 8-4. Rechazar una hipótesis nula cuando es cierta se denomina *error tipo I*, y su probabilidad (que, como hemos visto, es también el nivel de significancia de la prueba) se simboliza con α (*alfa*). Por otro lado, aceptar una hipótesis nula cuando es falsa se le llama *error tipo II*, y su probabilidad se simboliza con β (*beta*). Existe una relación entre estos dos tipos de errores: la probabilidad de cometer un tipo de error puede reducirse sólo si estamos dispuestos a aumentar la probabilidad de cometer el otro tipo de error. Observe en la gráfica (c) de la figura 8-4, que la región de aceptación es bastante pequeña (0.50 del área bajo la curva). Con una región de aceptación así de pequeña, rara vez aceptaremos una hipótesis nula cuando no sea cierta, pero como precio de esta segu-

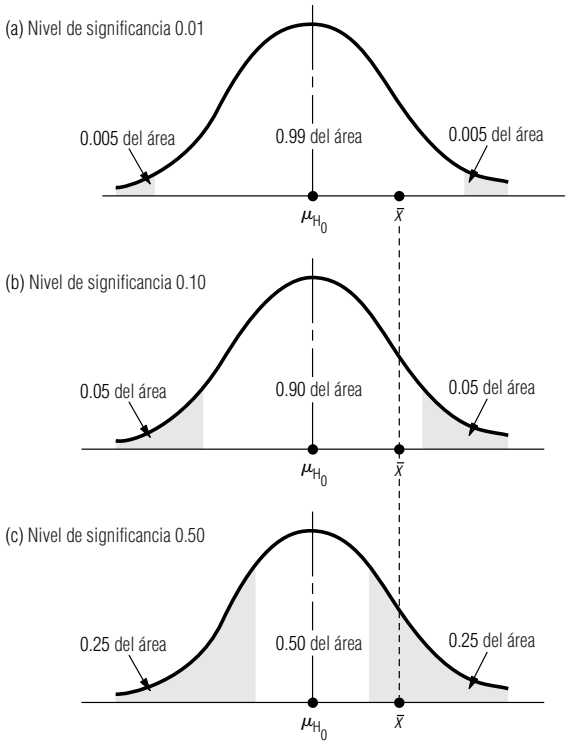


FIGURA 8-4

Tres niveles de significancia distintos

ridad, a menudo rechazaremos una hipótesis nula cuando sea cierta. Puesto de otra manera, con el fin de obtener una β baja, tendremos que tolerar una α alta. Para lidiar con este trueque o intercambio en situaciones personales y profesionales, los tomadores de decisiones deciden el nivel de significancia adecuado examinando los costos o la penalización vinculados con ambos tipos de error.

Preferencia por el error tipo I

Suponga que cometer el error tipo I (rechazar una hipótesis nula cuando es cierta) implica el tiempo y los problemas de volver a trabajar un lote de compuestos químicos que debieran haberse aceptado. Al mismo tiempo, cometer un error tipo II (aceptar una hipótesis nula cuando es falsa) significa arriesgarse a que todo un grupo de consumidores de este compuesto químico se envenenen. Evidentemente, la administración de esta compañía preferirá el error tipo I sobre el de tipo II y, como resultado, establecerá niveles de significancia muy altos en sus pruebas para obtener betas (β) bajas.

Preferencia por el error tipo II

Suponga, por otra parte, que cometer un error tipo I implica desarmar todo un motor en la fábrica, pero cometer un error tipo II implica que los distribuidores de automóviles realicen reparaciones de garantía relativamente baratas. Entonces es más probable que el fabricante prefiera el error tipo II y que establezca menores niveles de significancia en sus pruebas.

Decisión de qué tipo de distribución usar en la prueba de hipótesis

Selección de la distribución correcta antes de la prueba

Después de decidir el nivel de significancia a utilizar, nuestra siguiente tarea en la prueba de hipótesis consiste en determinar la distribución de probabilidad adecuada. Tenemos la opción de escoger entre la distribución normal, tabla 1 del apéndice, y la distribución *t*, tabla 2 del apéndice. Las reglas para elegir la distribución adecuada son similares a las que encontramos en el capítulo 7 sobre estimación. La tabla 8-1 resume las situaciones en las que es posible usar la distribución normal y la distribución *t* al hacer pruebas de medias. Más adelante, en este capítulo, examinaremos las distribuciones apropiadas para probar hipótesis sobre proporciones.

Uso del multiplicador de población finita

Recuerde otra regla que debe cumplirse al probar el valor hipotético de una media. Al igual que en la estimación, utilice el *multiplicador de población finita* siempre que la población sea finita en tamaño, el muestreo se haga sin reemplazo y la muestra sea de más del 5% de la población.

Pruebas de hipótesis de dos colas y de una cola

Descripción de una prueba de hipótesis de dos colas

En las pruebas de las siguientes medias hipotéticas de población, ilustraremos pruebas de dos colas y de una cola. Estos nuevos términos requieren explicación. Una *prueba de dos colas* rechaza la hipótesis nula si la media de muestra es significativamente mayor *o* menor que la media hipotética de la población. Por tanto, en una prueba de dos colas, existen *dos* regiones de rechazo. La figura 8-5 ilustra esto.

Una prueba de dos colas es apropiada cuando la hipótesis nula es $\mu = \mu_{H_0}$ (en donde μ_{H_0} es algún valor especificado) y la hipótesis alternativa es $\mu \neq \mu_{H_0}$. Suponga que un fabricante de focos desea producirlos con una vida media de $\mu = \mu_{H_0} = 1{,}000$ horas. Si el tiempo de vida es más corto, perderá clientes en favor de su competencia; si el tiempo de vida es más largo, tendrá un costo de producción muy alto porque los filamentos serán excesivamente gruesos. Para verificar que su proceso de producción sea adecuado, toma una muestra de la producción con el fin de probar la hipótesis H_0: $\mu = 1{,}000$. Como no quiere desviarse significativamente de 1,000 horas *en ninguna dirección*, la hipótesis alternativa adecuada es H_1: $\mu \neq 1{,}000$, y utiliza una prueba de dos colas. Esto es, rechaza la hipótesis nula si la vida media de los focos de la muestra está *muy por arriba* de 1,000 horas *o muy por abajo* de 1,000 horas.

En ocasiones debemos usar una prueba de una cola

Sin embargo, existen situaciones en las que no es apropiada una prueba de dos colas y debemos usar una prueba de una cola. Consideremos el caso de un mayorista que compra focos al fabricante del ejemplo anterior. El mayorista los compra en grandes lotes y no desea aceptar un lote de focos a menos que su vida media sea 1,000 horas o más. Cada vez que llega una remesa, el mayorista prueba una muestra para decidir si la acepta o no. La compañía rechazará el envío sólo si le parece que su vida media es menor que las 1,000 horas. Si cree que los focos son *mejores* que lo esperado (con una vida media superior a 1,000 horas), es claro que no rechazará la remesa, porque la vida más larga no tiene un costo adicional. Así que las hipótesis del mayorista son H_0: $\mu = 1{,}000$ horas y H_1: $\mu < 1{,}000$ horas. Rechaza H_0 sólo si la vida media de los focos muestreados es significativamente *menor* que 1,000 horas. La figura 8-6 ilustra esta situación. En ella podemos ver por qué esta prueba se conoce como *prueba de cola izquierda* (o *prueba de cola inferior*).

Pruebas de cola izquierda

En general, se utiliza una prueba de cola izquierda (de cola inferior) si las hipótesis son H_0: $\mu = \mu_{H_0}$ y H_1: $\mu < \mu_{H_0}$. En este caso, la evidencia muestral con la media de la muestra significativamente menor que la media hipotética de la población es la que nos lleva a rechazar la hipótesis nula en favor de la hipótesis alternativa. Dicho de otro modo, la región de rechazo está en la cola inferior (cola izquierdo) de la distribución de la media muestral, y ésa es la razón por la que la llamamos prueba de cola inferior.

Pruebas de cola derecha

La prueba de cola izquierda es uno de los dos tipos de pruebas de una cola. Como quizá habrá adivinado, el otro tipo de prueba de una cola es una *prueba de cola derecha* (o *prueba de cola superior*). Una prueba de cola superior se utiliza cuando las hipótesis son H_0: $\mu = \mu_{H_0}$ y H_1: $\mu > \mu_{H_0}$. Sólo los valores de la media de la muestra que son *significativamente mayores* que la media hipotética de la población harán que rechacemos la hipótesis nula en favor de la hipótesis alternativa. Esta prueba se denomina prueba de cola superior porque la región de rechazo está en la cola superior de la distribución de la media muestral.

Tabla 8-1

Condiciones para usar las distribuciones normal y *t* en la prueba de hipótesis sobre medias

	Cuando se conoce la desviación estándar de la población	Cuando no se conoce la desviación estándar de la población
El tamaño de muestra *n* es mayor que 30	Distribución normal, tabla *z*	Distribución normal, tabla *z*
El tamaño de muestra *n* es 30 o menos y suponemos que la población es normal o aproximadamente normal	Distribución normal, tabla *z*	Distribución *t*, tabla *t*

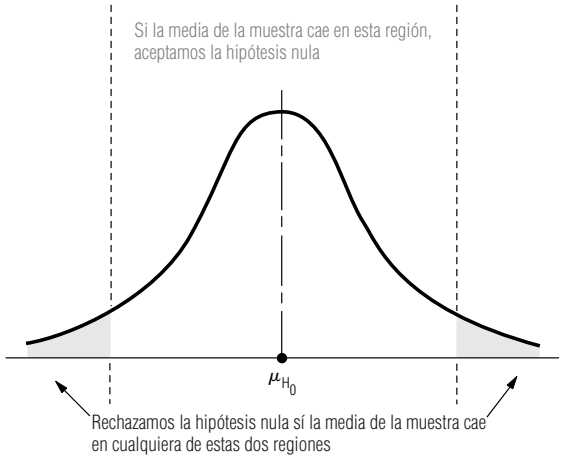


FIGURA 8-5

Prueba de hipótesis de dos colas, que muestra las dos regiones de rechazo

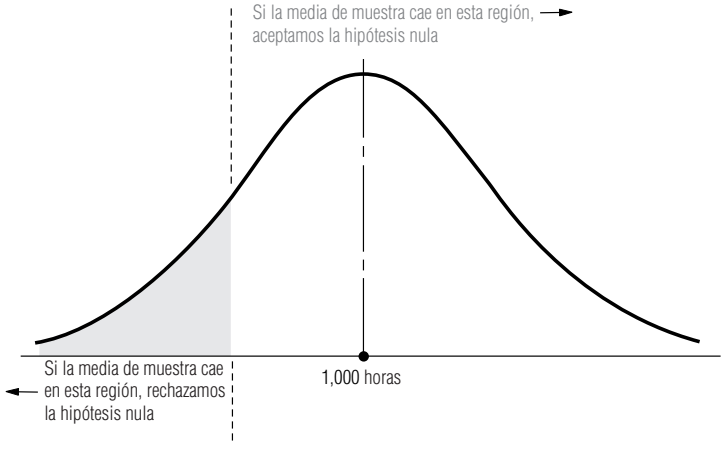


FIGURA 8-6

Pruebas de cola izquierda (prueba de cola inferior) con la región de rechazo en el lado izquierdo (lado inferior)

La figura 8-7 ilustra la siguiente situación; requiere el uso de una prueba de cola superior. Un administrador ha pedido a sus agentes de ventas que observen un límite en sus gastos de viaje. El administrador espera mantener los gastos en un promedio de \$100 diarios por vendedor. Un mes después de imponer el límite, se toma una muestra de gastos diarios entregados para ver si se ha observado el límite. La hipótesis nula es H_0: $\mu = \$100.00$, pero al administrador sólo le importan los gastos excesivamente altos. Por tanto, la hipótesis alternativa apropiada en este caso es H_1: $\mu > \$100.00$, y se utiliza una prueba de cola superior. La hipótesis nula se rechaza (y se toman medidas correctivas) sólo si la media de la muestra es significativamente mayor que \$100.00.

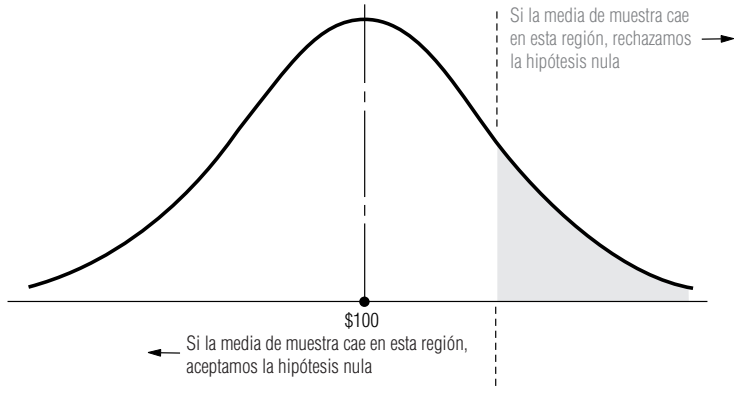


FIGURA 8-7

Prueba de cola derecha (de cola superior)

El aceptar H_0 no garantiza que H_0 sea cierta

Por último, debemos recordarle de nuevo, que en cada ejemplo de pruebas de hipótesis, cuando aceptamos una hipótesis nula con base en la información de la muestra, en realidad estamos diciendo que no hay evidencia estadística para rechazarla. No estamos afirmando que la hipótesis nula sea cierta. La única forma de *probar* una hipótesis nula es conocer el parámetro, y eso no es posible cuando se trata de muestreo. Así, aceptamos la hipótesis nula y nos comportamos como si fuera cierta, simplemente porque no podemos encontrar evidencia para rechazarla.

SUGERENCIAS Y SUPOSICIONES

Advertencia: no debe usar los resultados de la muestra para decidir entre usar una prueba de dos colas, de cola superior o de cola inferior. *Antes* de recolectar los datos, se determina la forma de la prueba a partir de lo que piensa o desea detectar el tomador de decisiones. Sugerencia: si los investigadores de mercados sospechan que las personas que compran cereal escarchado también compran *más* azúcar que quienes compran cereales sin endulzar, intentarán verificar su creencia sometiendo los datos a una prueba de *cola superior*. En caso de que la media de la muestra (sorprendentemente) fuera menor que el valor hipotético, esto no la convierte en una prueba de *cola inferior*, es sólo que los datos no apoyan la creencia original.

Ejercicios 8.3

Ejercicios de autoevaluación

EA 8-3 Para los siguientes casos, especifique qué distribución de probabilidad usar en una prueba de hipótesis:
a) H_0: $\mu = 27$, H_1: $\mu \neq 27$, $\bar{x} = 33$, $\hat{\sigma} = 4$, $n = 25$.
b) H_0: $\mu = 98.6$, H_1: $\mu > 98.6$, $\bar{x} = 99.1$, $\sigma = 1.5$, $n = 50$.
c) H_0: $\mu = 3.5$, H_1: $\mu < 3.5$, $\bar{x} = 2.8$, $\hat{\sigma} = 0.6$, $n = 18$.
d) H_0: $\mu = 382$, H_1: $\mu \neq 382$, $\bar{x} = 363$, $\sigma = 68$, $n = 12$.
e) H_0: $\mu = 57$, H_1: $\mu > 57$, $\bar{x} = 65$, $\hat{\sigma} = 12$, $n = 42$.

EA 8-4 Martha Inman, una ingeniera de seguridad en carreteras, decide probar la capacidad de carga de un puente que tiene 20 años. Dispone de una gran cantidad de datos de pruebas similares en el mismo tipo de puente. ¿Qué es adecuado, una prueba de una o de dos colas? Si la capacidad de carga mínima de este puente debe ser 10 toneladas, ¿cuáles son las hipótesis nula y alternativa?

Conceptos básicos

- **8-13** Formule las hipótesis nula y alternativa para probar si la nevada anual promedio en Buffalo, Nueva York, excede 45 pulg.
- **8-14** Describa qué representan normalmente las hipótesis nula y las alternativa en el proceso de prueba de hipótesis.
- **8-15** Defina el término *nivel de significancia*.
- **8-16** Defina los errores tipo I y tipo II.
- **8-17** En un juicio, la hipótesis nula es que un individuo es inocente de cierto crimen. ¿Preferirá el sistema legal cometer un error tipo I o un error tipo II con esta hipótesis?
- **8-18** ¿Cuál es la relación entre el nivel de significancia de una prueba y el error tipo I?
- **8-19** Si su objetivo es aceptar la hipótesis nula $\mu = 36.5$, con 96% de certeza cuando es cierta, y nuestro tamaño de muestra es 50, grafique las regiones de aceptación y rechazo para las siguientes hipótesis alternativas:
  a) $\mu \neq 36.5$.
  b) $\mu > 36.5$.
  c) $\mu < 36.5$.
- **8-20** Para los siguientes casos, especifique qué distribución de probabilidad debe usarse en una prueba de hipótesis:
  a) H_0: $\mu = 15$, H_1: $\mu \neq 15$, $\bar{x} = 14.8$, $\hat{\sigma} = 30$, $n = 35$.
  b) H_0: $\mu = 9.9$, H_1: $\mu \neq 9.9$, $\bar{x} = 10.6$, $\sigma = 2.3$, $n = 16$.

c) H_0: $\mu = 42$, H_1: $\mu > 42$, $\bar{x} = 44$, $\sigma = 4.0$, $n = 10$.
d) H_0: $\mu = 148$, H_1: $\mu > 148$, $\bar{x} = 152$, $\hat{\sigma} = 16.4$, $n = 29$.
e) H_0: $\mu = 8.6$, H_1: $\mu < 8.6$, $\bar{x} = 8.5$, $\hat{\sigma} = 0.15$, $n = 24$.

■ **8-21** Su hipótesis nula es que la batería de un marcapasos tiene una vida promedio de 300 días, con una hipótesis alternativa de que la vida de la batería es más de 300 días. Usted es ingeniero de control de calidad del fabricante de baterías.
a) ¿Preferiría cometer un error tipo I o un error tipo II?
b) Basándose en la respuesta al inciso anterior, ¿debería usar un nivel de significancia alto o bajo?

■ **8-22** ¿En qué condiciones es apropiado utilizar una prueba de una cola? ¿Y una de dos?

■ **8-23** Si usted decidió que una prueba de una cola es la apropiada, ¿cómo decide si debe ser una prueba de cola inferior o una de cola superior?

Aplicaciones

■ **8-24** El departamento de estadística instaló, el año pasado, luces, calentadores y aire acondicionado eficientes en el consumo de energía. Ahora desean determinar si el consumo promedio mensual de energía ha disminuido. ¿Deben realizar una prueba de una o de dos colas? Si el consumo promedio mensual anterior era 3,124 kilowatts hora, ¿cuáles son la hipótesis nula y alternativa?

■ **8-25** El doctor Ross Darrow opina que la nicotina de los cigarros ocasiona que los fumadores tengan ritmos cardiacos promedio más acelerados que los no fumadores. También piensa que los fumadores ansían la nicotina de los cigarros más que sólo fumar por la satisfacción física del acto y, en consecuencia, que el fumador promedio fumará más cigarros al día si cambia de una marca con un alto contenido de nicotina a otra con un contenido menor.
a) Suponga que Ross sabe que los no fumadores tienen un ritmo cardiaco de 78 pulsaciones por minuto. ¿Cuáles son las hipótesis nula y alternativa adecuadas para probar su primera opinión?
b) Durante los tres meses anteriores, ha estado observando una muestra de 48 individuos que fuman un promedio diario de 15 cigarros con alto contenido de nicotina. Les acaba de cambiar la marca de cigarros por otra con un bajo contenido de nicotina. Establezca las hipótesis nula y alternativa para probar su segunda opinión.

Soluciones a los ejercicios de autoevaluación

EA 8-3 a) *t* con 24 gl. b) Normal. c) *t* con 17 gl. d) Normal. e) *t* con 41 gl (entonces se usa una tabla normal).

EA 8-4 La ingeniera estará interesada en si un puente de esta edad puede soportar las capacidades mínimas de carga necesarias desde el punto de vista de seguridad. Entonces, quiere que su capacidad sea *mayor* que cierto nivel mínimo y usará una prueba de una cola (específicamente una prueba de cola superior o derecha). Las hipótesis son:

$$H_0: \mu = 10 \text{ toneladas} \qquad H_1: \mu > 10 \text{ toneladas}$$

8.4 Pruebas de hipótesis de medias cuando se conoce la desviación estándar de la población

Pruebas de dos colas de medias: prueba en la escala de la variable original

Un fabricante surte los ejes traseros para los camiones del Servicio Postal de Estados Unidos. Estos ejes deben soportar 80,000 libras por pulgada cuadrada en pruebas de carga, pero un eje excesivamente fuerte eleva los costos de producción de manera significativa. La larga experiencia indica que la desviación estándar de la resistencia de sus ejes es 4,000 libras por pulgada cuadrada. El fabrican-

te selecciona una muestra de 100 ejes de la producción, los prueba y encuentra que la capacidad de carga media de la muestra es 79,600 libras por pulgada cuadrada. Escritos en símbolos, los datos en este caso son

$$\mu_{H_0} = 80{,}000 \quad \leftarrow \text{Valores hipotetizados de la media de población}$$
$$\sigma = 4{,}000 \quad \leftarrow \text{Desviación estándar de la población}$$
$$n = 100 \quad \leftarrow \text{Tamaño de muestra}$$
$$\bar{x} = 79{,}600 \quad \leftarrow \text{Media de la muestra}$$

Cómo plantear el problema en símbolos

Si el fabricante de ejes utiliza un nivel de significancia (α) de 0.05 en la prueba, ¿satisfarán los ejes sus requerimientos de carga? En símbolos, podemos establecer el problema como:

$$H_0: \mu = 80{,}000 \quad \leftarrow \text{Hipótesis nula: la media real es 80,000 libras por pulgada cuadrada}$$
$$H_1: \mu \neq 80{,}000 \quad \leftarrow \text{Hipótesis alternativa: la media real no es 80,000 libras por pulgada cuadrada}$$
$$\alpha = 0.05 \quad \leftarrow \text{Nivel de significancia para probar esta hipótesis}$$

Cálculo del error estándar de la media

Como conocemos la desviación estándar de la población, y como el tamaño de la población es suficientemente grande para considerarlo infinito, podemos utilizar la distribución normal en nuestra prueba. Primero, calculamos el error estándar de la media usando la ecuación 6-1:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \qquad [6\text{-}1]$$
$$= \frac{4{,}000}{\sqrt{100}}$$
$$= \frac{4{,}000}{10}$$
$$= 400 \text{ libras por pulgada cuadrada} \quad \leftarrow \text{Error estándar de la media}$$

Ilustración del problema

Determinación de los límites de la región de aceptación

La figura 8-8 ilustra este problema, muestra el nivel de significancia de 0.05 como las dos regiones sombreadas que contienen, cada una, 0.025 del área. La región de aceptación de 0.95 contiene dos áreas iguales de 0.475 cada una. De la tabla de la distribución normal (tabla 1 del apéndice) podemos ver que el valor *z* apropiado para 0.475 del área bajo la curva es 1.96. Ahora podemos determinar los límites de la región de aceptación:

$$\mu_{H_0} + 1.96\sigma_{\bar{x}} = 80{,}000 + 1.96(400)$$
$$= 80{,}000 + 784$$
$$= 80{,}784 \text{ libras por pulgada cuadrada} \quad \leftarrow \text{Límite superior}$$

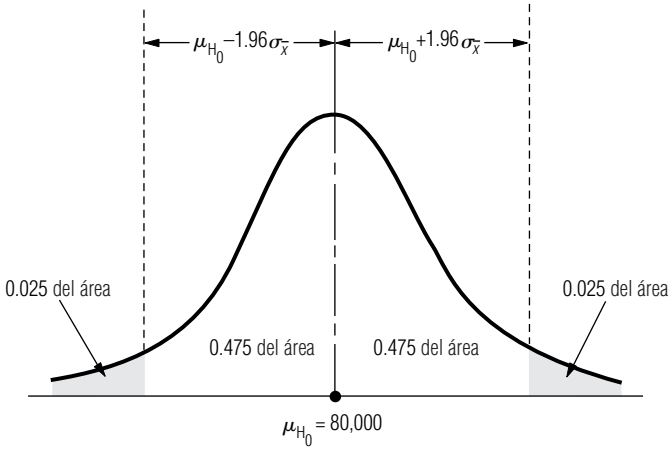


FIGURA 8-8

Prueba de hipótesis de dos colas al nivel de significancia de 0.05

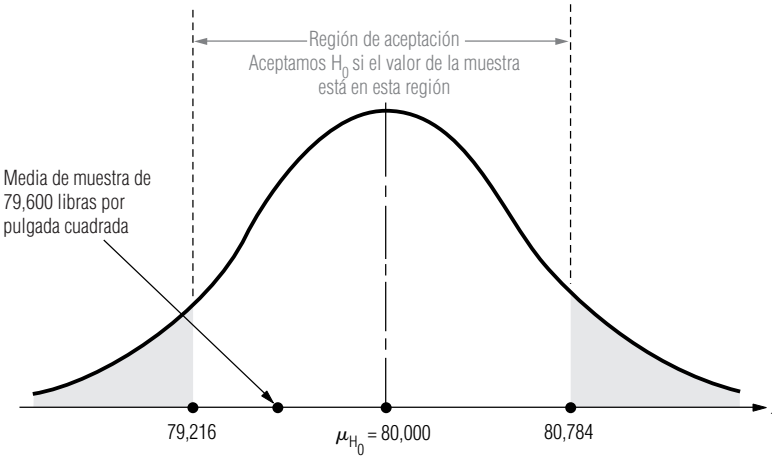


FIGURA 8-9

Prueba de hipótesis de dos colas al nivel de significancia de 0.05; muestra la región de aceptación y la media de la muestra

y

$$\begin{aligned}\mu_{H_0} - 1.96\sigma_{\bar{x}} &= 80{,}000 - 1.96(400)\\ &= 80{,}000 - 784\\ &= 79{,}216 \text{ libras por pulgada cuadrada} \leftarrow \text{Límite inferior}\end{aligned}$$

Interpretación de los resultados

Observe que hemos definido los límites de la región de aceptación (80,784 y 79,216) y la media de la muestra (79,600), y que se ilustran en la figura 8-9 en la escala de la variable original (libras por pulgada cuadrada). En la siguiente sección veremos otra forma de definir los límites de la región de aceptación y el valor de la media de la muestra. Evidentemente, la media de la muestra cae dentro de la región de aceptación; el fabricante debe aceptar la hipótesis nula porque no hay diferencia significativa entre la media hipotética de 80,000 y la media observada de los ejes de la muestra. Con base en esta muestra, el fabricante debe aceptar que la corrida de producción satisface los requerimientos de carga.

Prueba de hipótesis usando la escala estandarizada

En la prueba de hipótesis que acabamos de concluir se requirieron dos números para tomar la decisión: un *valor observado* calculado a partir de la muestra, y un *valor crítico* que define la frontera entre las regiones de aceptación y de rechazo. Veamos con cuidado cómo obtuvimos ese valor critico. Después de establecer el nivel de significancia de $\alpha = 0.05$, buscamos en la tabla 1 del apéndice, la distribución de probabilidad normal estándar, para encontrar que ± 1.96 son los valores z que dejaban 0.025 de probabilidad en cada extremo de la distribución.

Recordemos nuestro análisis de las variables normales estandarizadas en el capítulo 5. En vez de medir la variable en sus unidades originales, la variable estandarizada z nos dice a cuántas desviaciones estándar arriba ($z > 0$) o abajo ($z < 0$) de la media se encuentra nuestra observación. Entonces hay dos escalas de medición, la escala original o *sin procesar* y la *escala estandarizada*. La figura 8-10 es igual a la figura 8-9, pero incluye ambas escalas. Observe que nuestra media muestral de 79,600 libras por pulgada cuadrada está dada en la escala sin procesar, pero los valores críticos z de ± 1.96 están dados en la escala estandarizada. **Como estos dos números se dan en dos escalas distintas, no podemos compararlos directamente cuando probamos nuestras hipótesis. *Debemos convertir uno de ellos a la escala del otro*.**

Conversión del valor observado a la escala estandarizado

Hicimos nuestra prueba de hipótesis en la escala original al convertir los valores z críticos de ± 1.96 a los valores críticos de $\bar{x}$ en la escala original. Entonces, como el valor observado de $\bar{x}$ (79,600) cayó entre los límites inferior y superior de la región de aceptación (79,216 y 80,784), aceptamos la hipótesis nula. **En lugar de convertir los valores críticos z a la escala original, para ob-**

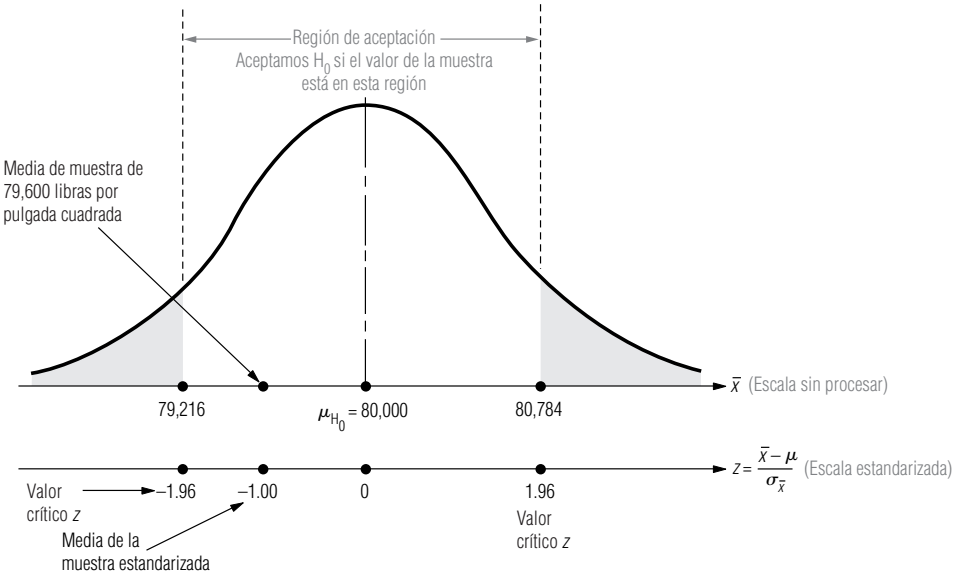


FIGURA 8-10

Prueba de hipótesis de dos colas al nivel de significancia de 0.05, que muestra la región de aceptación y la media de la muestra en las escalas sin procesar y estandarizada

tener números directamente comparables con el valor observado de $\bar{x}$, podríamos haber convertido nuestro valor observado de $\bar{x}$ a la escala estandarizada, utilizando la ecuación 6-2 para obtener un valor *z* observado, un número directamente comparable con los valores críticos *z*:

$$z = \frac{\bar{x} - \mu_{H_0}}{\sigma_{\bar{x}}}$$

El error estándar de la media de la ecuación 6-1

$$= \frac{79{,}600 - 80{,}000}{400}$$

La media de la muestra está a un error estándar abajo de la media de la población

$$= -1.00$$

La figura 8-10 también ilustra este valor observado en la escala estandarizada. Observe que el valor cae entre ± 1.96 de los límites inferior y superior de la región de aceptación de esta escala. Una vez más concluimos que se debe aceptar H_0: el fabricante debe aceptar que la corrida de producción reúne los requisitos de carga.

¿En qué difieren los dos métodos?

¿Cuál es la diferencia entre los dos métodos que acabamos de utilizar para probar nuestras hipótesis? Sólo en que definimos las unidades (o escala de medición) de manera distinta en cada método. **Sin embargo, los dos métodos siempre llevarán a las mismas conclusiones.** Algunas personas se sienten mejor usando la escala de la variable original; otras prefieren la estandarizada, que acabamos de explicar. Los resultados de la mayoría de los paquetes estadísticos de cómputo usan la escala estandarizada. En lo que resta del capítulo y en el capítulo 9, haremos las pruebas usando la escala estandarizada. Otra sugerencia: utilice el método que le resulte más cómodo.

El proceso de cinco pasos para la prueba de hipótesis usando la escala estandarizada

La tabla 8-2 resume el proceso de cinco pasos que utilizaremos en el resto de este capítulo y en todo el capítulo 9 para probar hipótesis.

Prueba de una cola para las medias

Para una prueba de una cola para una media, suponga que un hospital usa grandes cantidades de dosis envasadas de un medicamento particular. La dosis individual de esta medicina tiene 100 cm^3 (100 cc).

Tabla 8-2	Paso	Acción
Resumen del proceso de cinco pasos	1.	Decida si ésta es una prueba de dos colas o de una. Establezca sus hipótesis. Seleccione un nivel de significancia apropiado para esta decisión.
	2.	Decida qué distribución (*t* o *z*) es la adecuada (vea la tabla 8-1) y encuentre el (los) *valor(es) crítico(s)* para el nivel de significancia elegido en la tabla adecuada.
	3.	Calcule el *error estándar* del estadístico muestral. Use el error estándar para convertir el valor observado del estadístico en un valor estandarizado.
	4.	Bosqueje la distribución y marque la posición del valor de la muestra estandarizado y del valor o valores críticos para la prueba.
	5.	Compare el valor del estadístico muestral estandarizado con el (los) valor(es) crítico(s) para esta prueba e interprete el resultado.

La acción del medicamento es tal que el cuerpo tolera dosis excesivas sin sufrir daño. Por otra parte, las dosis insuficientes no producen el efecto médico deseado e interfieren con el tratamiento del paciente. El hospital ha adquirido la cantidad de medicamento que necesita al mismo fabricante durante varios años y sabe que la desviación estándar de la población es 2 cc. El hospital inspecciona, aleatoriamente, 50 dosis, tomadas de un envío muy grande y encuentra que la media de estas dosis es 99.75 cc.

$$\mu_{H_0} = 100 \quad \leftarrow \text{Valores hipotéticos de la media de la población}$$
$$\sigma = 2 \quad \leftarrow \text{Desviación estándar de la población}$$
$$n = 50 \quad \leftarrow \text{Tamaño de la muestra}$$
$$x = 99.75 \quad \leftarrow \text{Media de la muestra}$$

Si el hospital establece un nivel de significancia de 0.10 y nos pregunta si las dosis de esta entrega son demasiado pequeñas, ¿cómo podemos hallar la respuesta?

Paso 1: Establezca sus hipótesis, tipo de prueba y nivel de significancia

Para empezar, podemos expresar el problema en símbolos:

$$H_0: \mu = 100 \quad \leftarrow \text{Hipótesis nula: la media de las dosis de la remesa es 100 cc}$$
$$H_1: \mu < 100 \quad \leftarrow \text{Hipótesis alternativa: la media es menor que 100 cc}$$
$$\alpha = 0.10 \quad \leftarrow \text{Nivel de significancia para probar esta hipótesis}$$

Paso 2: Elija la distribución apropiada y encuentre el valor crítico

Como conocemos la desviación estándar de la población y *n* es mayor que 30, podemos utilizar la distribución normal. De la tabla 1 del apéndice podemos determinar que el valor de *z* para el 40% del área bajo la curva es 1.28, de modo que el valor crítico para la prueba de *cola inferior* es –1.28.

El hospital desea saber si las dosis reales son de 100 cc o si, por el contrario, las dosis son demasiado pequeñas. El hospital debe determinar que las dosis contienen *más* de una cierta cantidad, o debe rechazar el envío. Ésta es una prueba de *cola izquierda*, que se ilustra en la figura 8-11. Observe que la región sombreada corresponde al nivel de significancia de 0.10. También note que la región de aceptación consta del 40% en el lado izquierdo de la distribución *y* todo el lado derecho (50%), para un área total del 90%.

Paso 3: Calcule el error estándar y estandarice el estadístico muestral

Ahora podemos calcular el error estándar de la media, utilizando la desviación estándar de la población que conocemos y la ecuación 6-1 (debido a que el tamaño de población es suficientemente grande para considerarla infinito):

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \qquad [6\text{-}1]$$
$$= \frac{2}{\sqrt{50}}$$
$$= \frac{2}{7.07}$$
$$= 0.2829 \text{ cc} \quad \leftarrow \text{Error estándar de la media}$$

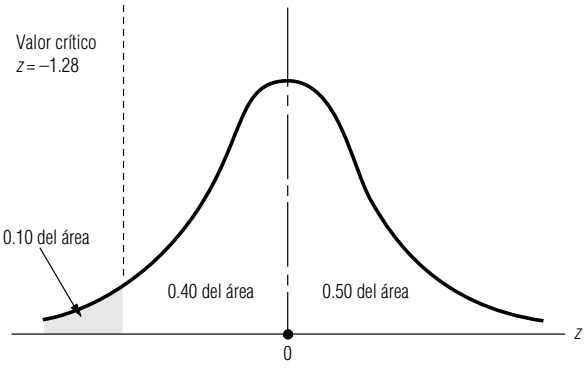


FIGURA 8-11

Prueba de hipótesis de cola izquierda a un nivel de significancia de 0.10

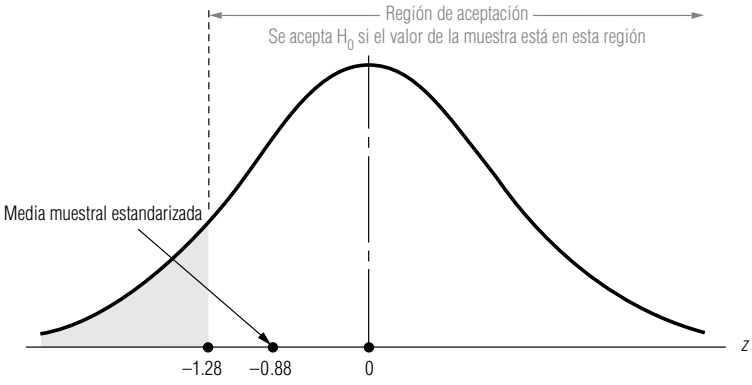


FIGURA 8-12

Prueba de hipótesis de cola izquierda al nivel de significancia de 0.10; muestra la región de aceptación y la media muestral estandarizada

Ahora usamos la ecuación 6-2 para *estandarizar* la media de la muestra, $\bar{x}$, restando μ_{H_0}, la media hipotética, y dividiendo entre $\sigma_{\bar{x}}$, el error estándar de la media:

$$z = \frac{\bar{x} - \mu_{H_0}}{\sigma_{\bar{x}}} \qquad [6\text{-}2]$$

$$= \frac{99.75 - 100}{0.2829}$$

$$= -0.88$$

Paso 4: Bosqueje la distribución y señale el valor de muestra y el valor crítico

Paso 5: Interprete el resultado

Al colocar el valor estandarizado en la escala *z* se observa que esta media muestral cae de lleno en la región de aceptación, como se muestra en la figura 8-12,

Por tanto, el hospital debe aceptar la hipótesis nula, porque la media observada de la muestra no es significativamente menor que la media hipotética de 100 cc. Con base en esta muestra de 50 dosis, el hospital debe concluir que las dosis de la entrega son suficientes.

SUGERENCIAS Y SUPOSICIONES

Existen muchas situaciones administrativas que requieren una prueba de una cola. Por ejemplo, un promotor de conciertos se interesa en atraer suficientes seguidores para *salir a mano o ganar*. Si llena el estadio y tiene que rechazar clientes, aumenta el prestigio del evento y no hay costo para él. Pero si no atrae suficientes asistentes puede tener problemas financieros. Él establecería una prueba de una cola, en palabras, "mayor o igual que 10,000 asistentes" (si 10,000 es su punto de equilibrio). Un distrito acuífero que está diseñando los límites de presión en su sistema de suministro tiene otra perspectiva. Al ingeniero hidráulico le interesa mantener la presión de agua *cerca de* cierto valor y usará una prueba de dos colas. Sugerencia: si la pregunta que debe responder se pone en palabra como *menor que*, *mayor que*, *menor o igual que*, o *mayor o igual que*, una prueba de una cola es apropiada. Si la pregunta se refiere a *diferente a*, *o cambio de* use una prueba de dos colas.

Ejercicios 8.4

Ejercicios de autoevaluación

EA 8-5 Hinton Press tiene la hipótesis de que la vida promedio de su prensa rotativa más grande es de 14,500 horas. Saben que la desviación estándar de la vida de una prensa es 2,100 horas. A partir de una muestra de 25 prensas, la compañía encuentra una media muestral de 13,000 horas. A un nivel de significancia de 0.01, ¿deben concluir que la vida promedio de las prensas es menor que las 14,500 horas hipotéticas?

EA 8-6 American Theaters sabe que cierta película de éxito se exhibió un promedio de 84 días en cada ciudad y que la desviación estándar correspondiente fue 10 días. El administrador del distrito sureste se interesó en comparar la popularidad de la película en su región con la que tuvo en otros cines de Estados Unidos. Eligió 75 salas al azar en su región y encontró que exhibieron la película un promedio de 81.5 días.

a) Establezca las hipótesis adecuadas para probar si hubo una diferencia significativa en la duración de la exhibición entre los teatros del sureste y el resto de Estados Unidos.
b) Pruebe estas hipótesis para un nivel de significancia del 1%.

Aplicaciones

8-26 Atlas Sporting Goods ha puesto en marcha una promoción especial para su estufa de propano y siente que la promoción debe provocar un cambio en el precio para el consumidor. Atlas sabe que antes de que comenzara la promoción, el precio promedio al menudeo de la estufa era $44.95, con una desviación estándar de $5.75. Atlas muestrea a 25 de sus minoristas después de iniciada la promoción y encuentra que el precio medio de las estufas es ahora $42.95. A un nivel de significancia de 0.02, ¿tiene Atlas razones para creer que el precio promedio al menudeo para el consumidor ha disminuido?

8-27 De 1980 a 1985, la tasa promedio de precios/utilidades (P/U) de los aproximadamente 1,800 valores inscritos en la Bolsa de Valores de Nueva York fue 14.35, con una desviación estándar de 9.73. En una muestra de 30 valores de la Bolsa, seleccionados al azar, la tasa P/U promedio en 1986 fue 11.77. ¿Esta muestra presenta evidencia suficiente para concluir (a un nivel de significancia de 0.05) que en 1986 la tasa P/U promedio para los valores de la Bolsa cambió su valor anterior?

8-28 Generally Electric ha desarrollado un nuevo foco cuyas especificaciones de diseño requieren una salida de luz de 960 lúmenes comparado con un modelo anterior que producía sólo 750 lúmenes. Los datos de la compañía indican que la desviación estándar de la salida de luz para este tipo de foco es 18.4 lúmenes. Para una muestra de 20 focos, el comité de pruebas encontró una salida de luz promedio de 954 lúmenes por foco. A un nivel de significancia de 0.05, ¿puede concluir Generally Electric que su nuevo foco produce la salida especificada de 950 lúmenes?

8-29 Maxwell's Hot Chocolate está preocupado por el efecto que la campaña de publicidad de café, de todo el año, pueda tener en las ventas de chocolate caliente. Las ventas semanales promedio de chocolate caliente hace dos años eran 984.7 libras y la desviación estándar era 72.6 libras. Maxwell seleccionó una muestra aleatoria de 30 semanas del año pasado y encontró ventas promedio de 912.1 libras.

a) Establezca las hipótesis adecuadas para probar si las ventas de chocolate han disminuido.
b) A un nivel de significancia del 2%, pruebe estas hipótesis.

8-30 La comisión promedio que cobran las compañías de corretaje de servicio completo en una venta de valores comunes es $144, con una desviación estándar de $52. Joel Freelander tomó una muestra aleatoria de 121 transacciones de sus clientes y determinó que habían pagado una comisión promedio de $151. A un nivel de significancia de 0.10, ¿puede concluir Joel que las comisiones de sus clientes son mayores que el promedio de la industria?

8-31 Históricamente, cada día, el Servicio de Aduanas de Estados Unidos intercepta alrededor de $28 millones en bienes de contrabando introducidos a ese país, con una desviación estándar de $16 millones al día. En 64 días de 1992, elegidos al azar, el Servicio de Aduanas interceptó un promedio de $30.3 millones en bienes de contrabando. ¿Indica esta muestra (a un nivel de significancia del 5%) que el Comisionado de Aduanas debería preocuparse por el incremento del contrabando por encima de su nivel histórico?

8-32 Antes del embargo petrolero de 1973 y los subsecuentes incrementos en el precio del petróleo crudo, el consumo de gasolina en Estados Unidos había aumentado a una tasa de ajuste estacional del 0.57% mensual, con una desviación estándar del 0.10% mensual. En 15 meses elegidos aleatoriamente entre 1975 y

1985, el consumo de gasolina tiene una tasa promedio de aumento de sólo 0.33% al mes. Para un nivel de significancia de 0.01, ¿puede concluir que el aumento en el uso de la gasolina se redujo como resultado del embargo y sus consecuencias?

■ **8-33** Bay City Bigleaguers, un equipo de béisbol semiprofesional, tiene al jugador líder de la liga en promedio de bateo durante muchos años. Durante los últimos años, el promedio de bateo de Joe Carver ha sido de .343, con una desviación estándar de .018. Sin embargo, este año, el promedio de Joe fue sólo .306. Joe está renegociando su contrato para el año siguiente, y el salario que podrá obtener depende en gran medida de su habilidad de convencer al dueño del equipo de que su promedio de bateo de este año no fue significativamente peor que en años anteriores. Si el dueño desea usar un nivel de significancia de 0.02, ¿reducirán el salario de Joe para el año próximo?

Soluciones a los ejercicios de autoevaluación

EA 8-5 $\sigma = 2{,}100 \quad n = 25 \quad \bar{x} = 13{,}000$

H_0: $\mu = 14{,}500 \quad H_1$: $\mu < 14{,}500 \quad \alpha = 0.01$

El límite inferior de la región de aceptación es $z = -2.33$, o

$$\bar{x} = \mu_{H_0} - z\sigma/\sqrt{n} = 14{,}500 - \frac{2.33(2{,}100)}{\sqrt{25}} = 13{,}521.4 \text{ horas}$$

Debido al valor z observado $= \dfrac{\bar{x} - \mu_{H_0}}{\sigma/\sqrt{n}} = \dfrac{13{,}000 - 14{,}500}{2{,}100/\sqrt{25}} = -3.57 < -2.33$

(es decir, $\bar{x} < 13{,}521.4$), debe rechazarse H_0. La vida promedio es significativamente menor que el valor hipotético.

EA 8-6 $\sigma = 10 \quad n = 75 \quad \bar{x} = 8.15$

H_0: $\mu = 84 \quad H_1$: $\mu \neq 84 \quad \alpha = 0.01$

Los límites de la región de aceptación son $z = \pm 2.58$ o

$$\bar{x} = \mu_{H_0} \pm z\sigma/\sqrt{n} = 84 \pm \frac{2.58(10)}{\sqrt{75}} = (81.02, 86.98) \text{ días}$$

Debido a que el valor z observado $= \dfrac{\bar{x} - \mu_{H_0}}{\sigma/\sqrt{n}} = \dfrac{81.5 - 84}{10/\sqrt{75}} = -2.17$, este valor y $\bar{x}$ están en la región de aceptación, de manera que no se rechaza H_0. La duración de la exhibición no es significativamente diferente de la de otras regiones.

8.5 Medición de la potencia de una prueba de hipótesis

¿Qué debe hacer una buena prueba de hipótesis?

Ahora que hemos considerado dos ejemplos de pruebas de hipótesis, resulta apropiado retroceder un poco para analizar lo que una buena prueba de hipótesis *debe hacer*. Idealmente, tanto α como β (las probabilidades de los errores tipo I y tipo II) deben ser pequeñas. Recuerde que un error tipo I se presenta cuando rechazamos una hipótesis nula que es cierta, y que α (el nivel de significancia de la prueba) *es* la probabilidad de cometer un error tipo I. En otras palabras, una vez que decidimos el nivel de significancia, no hay nada que podamos hacer respecto a α. Un error tipo II ocurre cuando aceptamos una hipótesis nula que es falsa; la probabilidad de un error tipo II es β. ¿Qué podemos decir respecto a β?

Significado de β y $1 - \beta$

Suponga que la hipótesis nula *es* falsa. Entonces los administradores desearían que la prueba de hipótesis la rechazara siempre. Desafortunadamente, las pruebas de hipótesis no pueden ser infalibles; algunas veces, cuando la hipótesis nula es falsa, una prueba no la rechaza y, por tanto, se comete un error tipo II. Cuando la hipótesis nula es falsa, μ (la media *verdadera* de la población) no

es igual a μ_{H_0} (la media *hipotética* de la población); μ es igual a algún otro valor. Por cada valor de μ para el que la hipótesis alternativa es cierta hay una probabilidad diferente (β) de aceptar incorrectamente la hipótesis nula. Claro que desearíamos que esta β (la probabilidad de aceptar una hipótesis nula cuando es falsa) fuera lo más pequeña posible o, de manera equivalente, nos gustaría que $1 - \beta$ (la probabilidad de rechazar una hipótesis nula cuando es falsa) fuera lo más grande posible.

Interpretación de los valores de $1 - \beta$

Puesto que rechazar una hipótesis nula cuando es falsa es justo lo que debe hacer una buena prueba, un valor alto de $1 - \beta$ (algo cerca de 1.0) significa que la prueba trabaja bastante bien (rechaza la hipótesis nula cuando es falsa); un valor bajo de $1 - \beta$ (cerca de 0.0) significa que la prueba trabaja muy mal (no rechaza la hipótesis nula cuando es falsa). Como el valor de $1 - \beta$ es la medida de qué tan bien trabaja la prueba, recibe el nombre de *potencia de la prueba.* Si graficamos los valores de $1 - \beta$ para cada valor de μ para el que la hipótesis alternativa es cierta, la curva resultante se conoce como *curva de potencia*.

Cálculo de los valores de $1 - \beta$

La gráfica (a) de la figura 8-13 reproduce la prueba de cola izquierda de la figura 8-11, pero con la escala sin procesar. La gráfica (b) de la figura 8-13 muestra la curva de potencia asociada con esta prueba. No es difícil calcular los valores $1 - \beta$ para graficar la curva de potencia; tres de estos puntos se muestran en la gráfica (b) de la figura 8-13. Recuerde que con esta prueba intentábamos decidir si se aceptar o no una entrega de medicamentos. Nuestra prueba dictaba que debíamos rechazar la hipótesis nula si la media estandarizada de la muestra era menor que -1.28, esto es, si la dosis media de la muestra era menor que $100.00 - 1.28\ (0.2829)$, o 99.64 cc.

Considere el punto *C* de la curva de potencia presentada en la gráfica (b) de la figura 8-13. La dosis media de la población es 99.42 cc; dado este valor, debemos calcular la probabilidad de que la media de una muestra aleatoria de 50 dosis de esta población sea menor que 99.64 cc (el punto abajo del cual decidimos rechazar la hipótesis nula). Ahora considere la gráfica (c) de la figura 8-13. Antes calculamos el error estándar de la media como 0.2829 cc, así que 99.64 cc está a $(99.64 - 99.42)/0.2829$, o 0.78 errores estándar arriba de 99.42 cc. Utilizando la tabla 1 del apéndice, podemos ver que la probabilidad de observar una media muestral menor que 99.64 cc y, por consiguiente, rechazar la hipótesis nula, es 0.7823, el área sombreada de la gráfica (c). Entonces, la potencia de la prue-

Interpretación de un punto en la curva de potencia

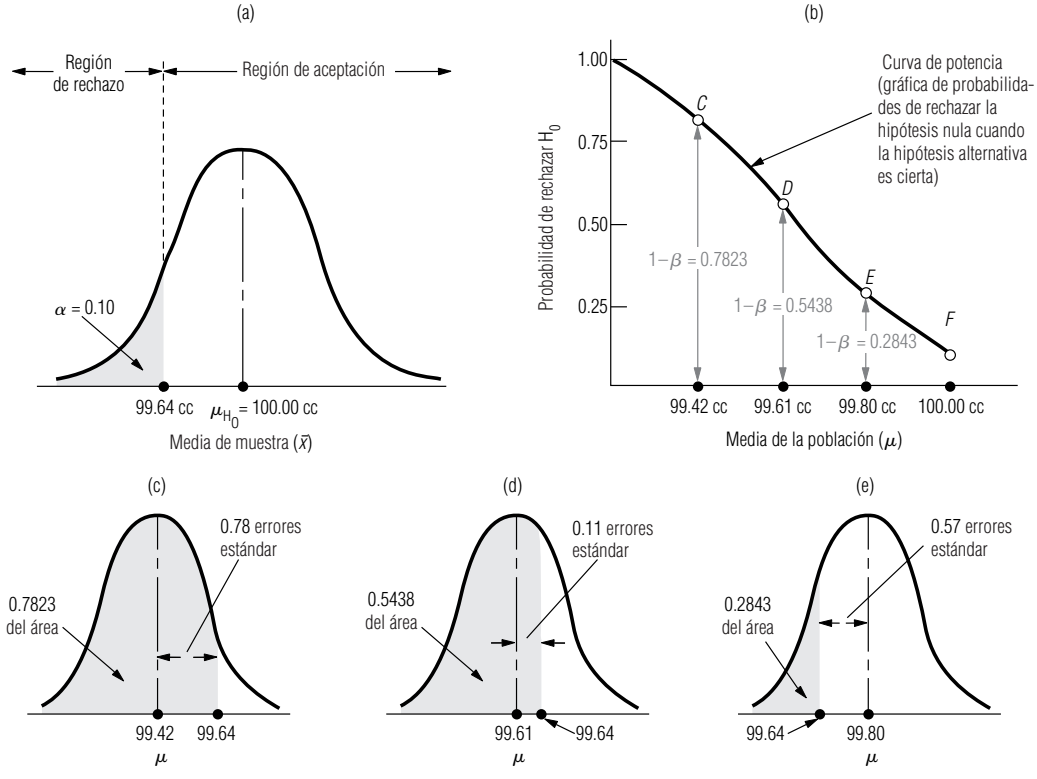


FIGURA 8-13

Prueba de hipótesis de cola izquierda, curva de potencia asociada y tres valores de μ

ba ($1 - \beta$) en $\mu = 99.42$ es 0.7823. Esto sólo significa que si $\mu = 99.42$, la probabilidad de que esta prueba rechace la hipótesis nula cuando es falsa es 0.7823.

Ahora observe el punto *D* de la gráfica (b) de la figura 8-13. Para esta dosis media de la población de 99.61 cc, ¿cuál es la probabilidad de que la media de una muestra aleatoria de 50 dosis de esta población sea menor que 99.64 cc y que esto ocasione que la prueba rechace la hipótesis nula? Vea la gráfica (d) de la figura 8-13. En ella 99.64 cc está a (99.64 − 99.61)/0.2829, o 0.11 errores estándar arriba de 99.61 cc. Utilizando de nuevo la tabla 1 del apéndice, podemos ver que la probabilidad de observar una media muestral menor que 99.64 y por consiguiente de rechazar la hipótesis nula es 0.5438, el área sombreada de la gráfica (d) de la figura 8-13. Por tanto, la potencia de la prueba ($1 - \beta$) para $\mu = 99.61$ cc es 0.5438.

Punto terminal de la curva de potencia

Con el mismo procedimiento para el punto *E*, encontramos que la potencia de la prueba en $\mu = 99.80$ cc es 0.2843; esto se ilustra como el área sombreada de la gráfica (e) de la figura 8-13. Los valores de $1 - \beta$ continúan disminuyendo a la derecha del punto *E*. ¿Qué tan abajo llegan? Al acercarse cada vez más la media de población a 100.00 cc, la potencia de la prueba ($1 - \beta$) debe acercarse cada vez más a la probabilidad de rechazar la hipótesis nula cuando la media de población es exactamente 100.00 cc, y sabemos que *esa* probabilidad es nada menos que el nivel de significancia de la prueba, en este caso, 0.10. Así, la curva termina en el punto *F*, que está a una altura de 0.10 directamente arriba de la media de la población.

Interpretación de la curva de potencia

¿Qué nos dice la curva de potencia de la gráfica (b) de la figura 8-13? Sólo que cuando la entrega es menos satisfactoria (al disminuir las dosis del envío), nuestra prueba es más poderosa (tiene una mayor probabilidad de reconocer que la entrega es insatisfactoria). Pero también nos muestra que debido al error de muestreo, cuando la dosis es sólo ligeramente menor que 100.00 cc, la potencia de la prueba para reconocer esta situación es bastante baja. Así, si resultara completamente insatisfactorio tener *cualquier* dosis con menos de 100.00 cc, la prueba analizada no sería apropiada.

SUGERENCIAS Y SUPOSICIONES

Por supuesto, siempre es mejor usar las pruebas de hipótesis con la mayor potencia. Pero también se sabe que cierta proporción del tiempo, todas las pruebas de hipótesis fallan y no rechazan la hipótesis nula cuando es falsa o no la aceptan cuando es cierta (éste es el lenguaje estadístico para decir que cuando una prueba falla, nos convencerá de que las cosas no han cambiado cuando de hecho cambiaron, o nos convencerá de que cambiaron cuando en realidad no lo hicieron). Ése es el precio que se paga al usar muestreo en las pruebas de hipótesis. El fracaso de una prueba para obtener el resultado correcto se debe a errores de muestreo. La única manera de evitar tales errores es examinar toda la población y eso es físicamente imposible, o bien, demasiado costoso.

Ejercicios 8.5

Ejercicios de autoevaluación

EA 8-7 Vea el ejercicio 8-32. Calcule la potencia de la prueba para $\mu = 0.50$, 0.45 y 0.40% por mes.

EA 8-8 En el ejercicio 8-32, ¿qué ocurre con la potencia de la prueba para $\mu = 0.50$, 0.45 y 0.40% por mes si se cambia el nivel de significancia a 0.04?

Aplicaciones

- **8-34** Vea el ejercicio 8-31. Calcule la potencia de la prueba para $\mu = \$28$, \$29 y \$30 millones.
- **8-35** Vea el ejercicio 8-30. Calcule la potencia de la prueba para $\mu = \$140$, \$160 y \$175.
- **8-36** En el ejercicio 8-31, ¿qué sucede con la potencia de la prueba para $\mu = \$28$, \$29 y \$30 millones si el nivel de significancia se cambia a 0.02?
- **8-37** En el ejercicio 8-30, ¿qué le ocurre a la potencia de la prueba para $\mu = \$140$, \$160 y \$175 al mes si el nivel de significancia se cambia a 0.05?

Soluciones a los ejercicios de autoevaluación

EA 8-7 Del ejercicio 8-32, se tiene $\sigma = 0.10$, $n = 15$, H_0: $\mu = 0.57$, H_1: $\mu < 0.57$.
Para $\alpha = 0.01$, el límite inferior de la región de aceptación es

$$\mu_{H_0} - 2.33\sigma/\sqrt{n} = 0.57 - 2.33(0.10)/\sqrt{15} = 0.510$$

a) Para $\mu = 0.50$, la potencia de la prueba es

$$P(\bar{x} < 0.510) = P(z < \frac{0.510 - 0.50)}{0.10/\sqrt{15}} = P(z < 0.39) = 0.5 + 0.1517 = 0.6517$$

b) Para $\mu = 0.45$, la potencia de la prueba es

$$P(\bar{x} < 0.510) = P(z < \frac{0.510 - 0.45)}{0.10/\sqrt{15}} = P(z < 2.32) = 0.5 + 0.4898 = 0.9898$$

c) Para $\mu = 0.40$, la potencia de la prueba es

$$P(\bar{x} < 0.510) = P(z < \frac{0.510 - 0.40)}{0.10/\sqrt{15}} = P(z < 4.26) = 1.0000$$

EA 8-8 Para $\alpha = 0.04$, el límite inferior de la región de aceptación es

$$\mu_{H_0} - 1.75\sigma/\sqrt{n} = 0.57 - 1.75(0.10)/\sqrt{15} = 0.525$$

a) Para $\mu = 0.50$, la potencia de la prueba es

$$P(\bar{x} < 0.525) = P(z < \frac{0.525 - 0.50)}{0.10/\sqrt{15}} = P(z < 0.97) = 0.5 + 0.3340 = 0.8340$$

b) Para $\mu = 0.45$, la potencia de la prueba es

$$P(\bar{x} < 0.525) = P(z < \frac{0.525 - 0.45)}{0.10/\sqrt{15}} = P(z < 2.90) = 0.5 + 0.4981 = 0.9981$$

c) Para $\mu = 0.40$, la potencia de la prueba es

$$P(\bar{x} < 0.525) = P(z < \frac{0.525 - 0.40)}{0.10/\sqrt{15}} = P(z < 4.84) = 1.0000$$

8.6 Prueba de hipótesis para proporciones: muestras grandes

Pruebas de dos colas para proporciones

Manejo de proporciones

En esta sección, lo aprendido respecto a las pruebas de medias, se aplicará a las pruebas para *proporciones* (esto es, la proporción de ocurrencias en una población). Pero antes revisaremos las importantes conclusiones a las que llegamos en el capítulo 7. Primero, recordemos que la binomial es la distribución teóricamente correcta para usarse al trabajar con proporciones, porque los datos son discretos, no continuos. Al aumentar el tamaño de la muestra, la distribución binomial se aproxima a la normal en sus características y podemos utilizar la distribución normal para aproximar la distribución de muestreo. Específicamente, *np y nq cada una debe ser al menos 5* para poder utilizar la distribución normal como aproximación de la binomial.

Imagine, por ejemplo, una compañía que está evaluando a cuáles de sus empleados ascender, determinando la proporción de aquellos cuya capacidad, capacitación y experiencia de supervisión los califican para pasar al siguiente nivel administrativo. El director de recursos humanos dice al presidente que aproximadamente el 80%, o 0.8, de los empleados de la compañía son "aptos para un as-

censo". El presidente reúne un comité especial para evaluar la capacidad de ascenso de todos los empleados. Este comité hace entrevistas a fondo con 150 empleados y encuentra que, a su juicio, sólo el 70% de la muestra está calificada para el ascenso.

$p_{H_0} = 0.8$ ← Valor hipotético de la proporción de éxitos de la población (considerados aptos para el ascenso, en este caso)

$q_{H_0} = 0.2$ ← Valor hipotético de la proporción de fracasos de la población (considerados no aptos para el ascenso)

$n = 150$ ← Tamaño de muestra

$\overline{p} = 0.7$ ← Proporción apta para el ascenso de la muestra

$\overline{q} = 0.3$ ← Proporción no apta para el ascenso de la muestra

Paso 1: Establezca sus hipótesis, tipo de prueba y nivel de significancia

El presidente desea probar al nivel de significancia de 0.05 la hipótesis de que 0.8 de los empleados son aptos para el ascenso:

$H_0: p = 0.8$ ← Hipótesis nula: 80% de los empleados es apto

$H_1: p \neq 0.8$ ← Hipótesis alternativa: la proporción de empleados aptos no es el 80%

$\alpha = 0.05$ ← Nivel de significancia

Paso 2: Elija la distribución apropiada y encuentre el valor *crítico*

En este caso, la compañía desea saber si la proporción verdadera es mayor o menor que la proporción hipotética. Por consiguiente, resulta adecuado efectuar una prueba de dos colas para una proporción; la figura 8-14 muestra la gráfica. El nivel de significancia corresponde a las dos regiones sombreadas, cada una con 0.025 del área. La región de aceptación de 0.95 se ilustra como dos áreas de 0.475 cada una. Puesto que *np* y *nq* son cada una mayores que 5, podemos utilizar la aproximación normal de la distribución binomial. De la tabla 1 del apéndice, podemos determinar que el valor crítico de *z* para 0.475 del área bajo la curva es 1.96.

Paso 3: Calcule el error estándar y estandarice el estadístico de la muestra

Podemos calcular el error estándar de la proporción, utilizando los valores hipotéticos de p_{H_0} y q_{H_0} en la ecuación 7-4:

$$\sigma_{\overline{p}} = \sqrt{\frac{p_{H_0} q_{H_0}}{n}} \qquad [7\text{-}4]$$

$$= \sqrt{\frac{(0.8)(0.2)}{150}}$$

$$= \sqrt{0.0010666}$$

$$= 0.0327 \leftarrow \text{Error estándar de la proporción}$$

A continuación estandarizamos la proporción de la muestra dividiendo la diferencia entre la proporción de la muestra observada $\overline{p}$ y la proporción hipotética, p_{H_0}, entre el error estándar de la proporción.

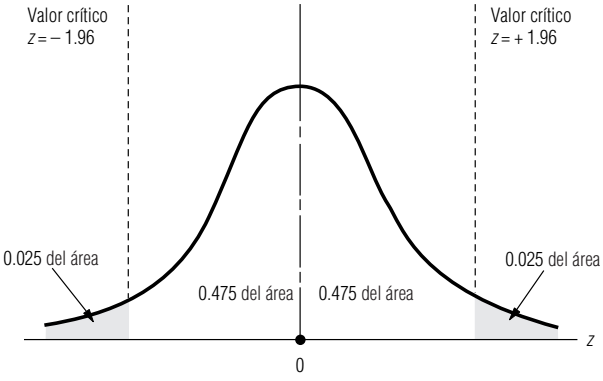


FIGURA 8-14

Prueba de hipótesis de dos colas para una proporción al nivel de significancia de 0.05

$$z = \frac{\bar{p} - p_{H_0}}{\sigma_{\bar{p}}}$$

$$= \frac{0.7 - 0.8}{0.0327}$$

$$= -3.06$$

Paso 4: Bosqueje la distribución y señale el valor de la muestra y los valores críticos

Paso 5: Interprete el resultado

Al señalar la proporción de la muestra estandarizada que se calculó, -3.06, en un bosquejo de la distribución muestral, es claro que esta muestra cae fuera de la región de aceptación, como se ve en la figura 8-15.

Entonces, en este caso, el presidente debe rechazar la hipótesis nula y concluir que *existe* una diferencia significativa entre la proporción hipotética de empleados aptos para ascenso (0.8) del director de recursos humanos y la proporción observada de empleados aptos en la muestra. De esto, debe inferir que la proporción real de empleados aptos en toda la compañía no es el 80%.

Pruebas de una cola para proporciones

Una prueba de una cola para una proporción es conceptualmente equivalente a una prueba de una cola para una media, como puede ilustrarse con este ejemplo. Un miembro de un grupo de interés público preocupado por la contaminación ambiental afirma, en una audiencia pública, que "menos del 60% de las plantas industriales de esta área cumple con los estándares de cuidado del ambiente". A la reunión asistió una funcionaria de 1a Agencia de Protección Ambiental (APA) quien cree que el 60% de las plantas sí cumple con los estándares; decide probar esa hipótesis al nivel de significancia de 0.02.

Paso 1: Establezca sus hipótesis, tipo de prueba y nivel de significancia

H_0: $p = 0.6$ ← Hipótesis nula: la proporción de plantas que cumplen con los estándares de contaminación ambiental es 0.6

H_1: $p < 0.6$ ← Hipótesis alternativa: la proporción que cumple con los estándares de contaminación ambiental es menor que 0.6

$\alpha = 0.02$ ← Nivel de significancia para probar la hipótesis

La funcionaria realiza una investigación completa de los registros de su oficina. Muestrea 60 plantas de una población de más de 10,000 y encuentra que 33 cumple con los estándares de cuidado del ambiente. ¿Es válida la afirmación del miembro del grupo de interés público?

Comencemos por resumir el caso con símbolos:

$p_{H_0} = 0.6$ ← Valor hipotético de la proporción de población que cumple con los estándares de contaminación ambiental

$q_{H_0} = 0.4$ ← Valor hipotético de la proporción de población que no cumple y que, por tanto, está contaminando

$n = 60$ ← Tamaño de la muestra

$\bar{p} = 33/60$ o 0.55 ← Proporción de muestra que no contamina

$\bar{q} = 27/60$ o 0.45 ← Proporción de muestra que contamina

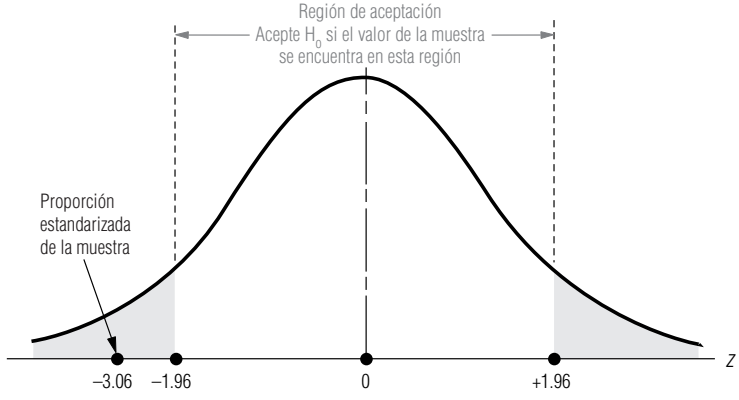


FIGURA 8-15

Prueba de hipótesis de dos colas para una proporción al nivel de significancia de 0.05; indica la región de aceptación y la proporción estandarizada de la muestra

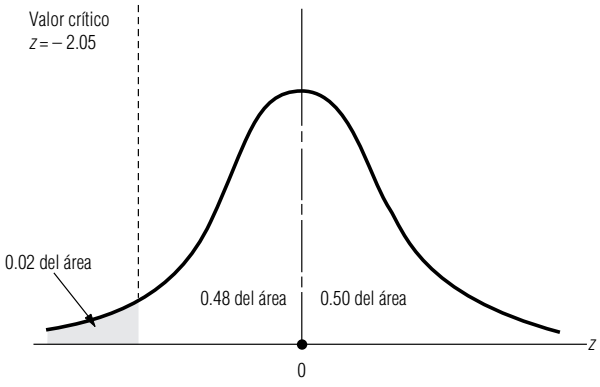


FIGURA 8-16

Prueba de hipótesis de una cola para una proporción al nivel de significancia de 0.02

Ésta es una prueba de una cola: la funcionaria de APA se pregunta sólo si la proporción real es menor que 0.6. Específicamente, ésta es una prueba de cola izquierda. Para rechazar la hipótesis nula que establece que la proporción verdadera de plantas que cumplen es el 60%, la representante de APA debe aceptar la hipótesis alternativa que dice que menos de 0.6 cumplen. La figura 8-16 ilustra esta prueba de hipótesis.

Paso 2: Elija la distribución apropiada y encuentre el valor crítico

Como *np* y *nq* son cada uno mayores que 5, podemos usar la aproximación normal de la distribución binomial. El valor crítico de *z* de la tabla 1 del apéndice para 0.48 del área bajo la curva es 2.05.

Paso 3: Calcule el error estándar y estandarice el estadístico de la muestra

A continuación, podemos calcular el error estándar de la proporción usando la proporción hipotética de la población de la siguiente manera:

$$\sigma_{\bar{p}} = \sqrt{\frac{p_{H_0} q_{H_0}}{n}} \qquad [7\text{-}4]$$

$$= \sqrt{\frac{(0.6)(0.4)}{60}}$$

$$= \sqrt{0.004}$$

$$= 0.0632 \leftarrow \text{Error estándar de la proporción}$$

Y estandarizamos la proporción de la muestra dividiendo la diferencia entre la población de la muestra observada, $\bar{p}$, y la proporción hipotética, p_{H_0}, entre el error estándar de la proporción.

$$z = \frac{\bar{p} - p_{H_0}}{\sigma_{\bar{p}}}$$

$$= \frac{0.55 - 0.6}{0.0632}$$

$$= -10.79$$

Paso 4: Bosqueje la distribución y señale el valor de la muestra y los valores críticos

La figura 8-17 ilustra dónde se encuentra la proporción de la muestra en relación con el valor crítico, −2.05. Al observar esta figura, podemos ver que la proporción de la muestra cae dentro de la región de aceptación. Entonces, la funcionaria de APA debe aceptar la hipótesis nula de que la proporción real de plantas que cumplen es 0.6. **Aunque la proporción observada de la muestra es menor que 0.6, *no es significativamente menor que* 0.6, es decir, no está tan abajo de 0.6 para aceptar la afirmación del miembro del grupo de interés público.**

Paso 5: Interprete el resultado

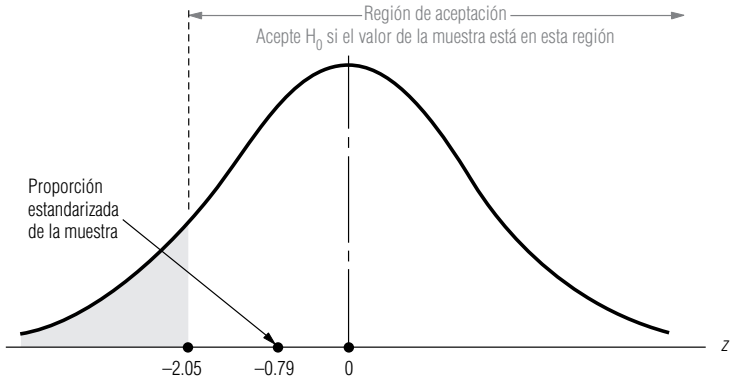


FIGURA 8-17

Prueba de hipótesis de una cola (cola izquierda) al nivel de significancia de 0.02; indica la región de aceptación y la proporción estandarizada de la muestra

SUGERENCIAS Y SUPOSICIONES

Advertencia: cuando se realizan pruebas de hipótesis que involucran proporciones, se usa la distribución binomial como la distribución muestral, a menos que *np* y *nq* sean *ambos* al menos 5. En ese caso se puede usar la distribución normal como una aproximación de la binomial sin problemas. Por fortuna, en la práctica, las pruebas de hipótesis para proporciones casi siempre usan muestras suficientemente grandes, de manera que esta condición se cumple. Aun cuando no sea así, no es tan difícil la aritmética de la distribución binomial o el uso de la tabla binomial.

Ejercicios 8.6

Ejercicios de autoevaluación

EA 8-9 Un fabricante de salsa de tomate está en proceso de decidir si produce o no una versión picante. El departamento de investigación de mercados de la compañía usó un sondeo telefónico nacional de 6,000 casas y encontró que 335 de ellos comprarían la salsa con picante. Un estudio más extenso hecho dos años antes mostró que el 5% de las casas compraría la marca. Con un nivel de significancia del 2%, ¿debe la compañía concluir que hay un incremento en el interés por el sabor con picante?

EA 8-10 Steve Cutter desea comparar la confiabilidad de las podadoras Big Blade que vende en su ferretería con la de las vendidas por la marca en todo el país. Steve sabe que sólo el 15% de todas las podadoras Big Blade necesitan reparaciones durante el primer año. Una muestra de 120 de los clientes de Steve reveló que exactamente 22 de ellos requirieron reparaciones en el primer año. Con un nivel de significancia de 0.02, ¿existe evidencia de que la confiabilidad de las podadoras Big Blade que vende Steve difiera de las que se venden en el todo el país?

Aplicaciones

■ **8-38** Grant, Inc., un fabricante de blusas de vestir para mujer, sabe que su marca se vende en 19% de las tiendas de ropa para mujer ubicadas al este del río Mississippi. Grant muestreó recientemente 85 tiendas de ropa para mujer en la ribera oeste del río y encontró que 14.12% de las tiendas vendía la marca. A un nivel de significancia del 0.04, ¿existe evidencia de que Grant tenga peor distribución en la ribera oeste que en la este del Mississippi?

■ **8-39** De un total de 10,200 préstamos otorgados por una unión de crédito de empleados del Estado en el último periodo de cinco años, se muestrearon 350 para determinar qué proporción de los préstamos se otorgaron a mujeres. Esta muestra indicó que el 39% de los créditos fue dado a empleadas. Un censo completo de préstamos de hace cinco años mostraba que el 41% de los prestatarios eran mujeres. A un nivel de significancia del 0.02, ¿puede concluir que la proporción de préstamos otorgados a mujeres ha cambiado significativamente en los últimos cinco años?

■ **8-40** Los laboratorios Feronetics se especializan en el uso de técnicas de reproducción de genes para lograr nuevos compuestos farmacéuticos. Recientemente desarrolló un atomizador nasal que contiene interferón, con el que se cree habrá de limitarse la transmisión del catarro común en las familias. En la población general, a 15.1% de todos los individuos les dará catarro ocasionado por un rinovirus una vez que otro miembro de la familia lo haya contraído. El atomizador de interferón fue probado en 180 personas, en cuyas familias uno de los miembros contrajo posteriormente un catarro ocasionado por un rinovirus. Sólo 17 de los sujetos de la prueba desarrollaron catarros similares.

a) A un nivel de significancia de 0.05, ¿debería concluir Feronetics que el nuevo atomizador efectivamente reduce la transmisión de catarros?

b) ¿Qué debería concluir en el nivel $\alpha = 0.02$?

c) Con base en estos resultados, ¿cree que se le debería permitir a Feronetics comercializar el nuevo atomizador? Explique su respuesta.

■ **8-41** Algunos teóricos financieros creen que los precios diarios del mercado de valores constituyen una "caminata aleatoria con tendencia positiva". Si esto es correcto, entonces el promedio industrial Dow Jones debería mostrar una ganancia en más del 50% de todos los días de actividad financiera. Si el promedio se incrementó en 101 de 175 días escogidos aleatoriamente, ¿qué piensa de la teoría sugerida? Use un nivel de significancia de 0.01.

■ **8-42** MacroSwift estimó el año pasado que el 35% de los compradores potenciales de software planeaba esperar hasta que se liberara una actualización de Window Panes para comprar el nuevo sistema operativo. Después de una campaña publicitaria para dar confianza al público, MacroSwift encuestó a 3,000 personas y encontró que 950 todavía se mostraban renuentes. Con el 5% de nivel de significancia, ¿puede la compañía concluir que la proporción de personas renuentes ha disminuido?

■ **8-43** Rick Douglas, el nuevo gerente de Food Barn, está interesado en el porcentaje de clientes totalmente satisfechos con la tienda. El gerente anterior tenía el 86% de clientes totalmente satisfechos y Rick asegura que lo mismo se cumple hoy. Rick obtuvo una muestra de 187 clientes y encontró que 157 estaban satisfechos por completo. Con un nivel de significancia del 1%, ¿existe evidencia de que la afirmación de Rick es válida?

Soluciones a los ejercicios de autoevaluación

EA 8-9 $n = 6{,}000 \quad \bar{p} = 335/6{,}000 = 0.05583$

$H_0: p = 0.05 \quad H_1: p > 0.05 \quad \alpha = 0.02$

El límite superior de la región de aceptación es $z = 2.05$, o

$$\bar{p} = p_{H_0} + z\sqrt{\frac{p_{H_0}q_{H_0}}{n}} = 0.05 + 2.05\sqrt{\frac{0.05(0.95)}{6{,}000}} = 0.05577$$

Debido a que el valor z observado $= \dfrac{\bar{p} - p_{H_0}}{\sqrt{p_{H_0}q_{H_0}/n}} = \dfrac{0.05583 - 0.05}{\sqrt{0.05(0.95)/6{,}000}} = 2.07 > 2.05$ (o $\bar{p} >$ 0.05577), debe rechazarse H_0 (pero por poco). El interés actual es significativamente mayor que el interés hace 2 años.

EA 8-10 $n = 120 \quad \bar{p} = 22/120 = 0.1833$

$H_0: p = 0.15 \quad H_1: p \neq 0.15 \quad \alpha = 0.02$

Los límites de la región de aceptación son $z = \pm 2.33$, o

$$\bar{p} = p_{H_0} \pm z\sqrt{\frac{p_{H_0}q_{H_0}}{n}} = 0.15 \pm 2.33\sqrt{\frac{0.15(0.85)}{120}} = (0.0741, 0.2259)$$

Debido a que el valor z observado $= \dfrac{\bar{p} - p_{H_0}}{\sqrt{p_{H_0}q_{H_0}/n}} = \dfrac{0.1833 - 0.15}{\sqrt{0.15(0.85)/120}} = 1.02 < 2.33$ (o $\bar{p} = 0.1833$, que está entre 0.0741 y 0.2259), no se rechaza H_0. Las podadoras de Steve no son significativamente diferentes en confiabilidad de las vendidas en todo el país.

8.7 Pruebas de hipótesis de medias cuando no se conoce la desviación estándar de la población

Cuándo usar la distribución *t*

Cuando estimamos los intervalos de confianza en el capítulo 7, aprendimos que la diferencia de tamaño entre muestras grandes y pequeñas es importante cuando no se conoce la desviación estándar de la población σ y es necesario estimarla a partir de la desviación estándar de la muestra. Si el tamaño de la muestra n es 30 o menos y σ se desconoce, debemos utilizar la distribución *t.* La distribución *t* apropiada tiene $n - 1$ grados de libertad. Estas reglas también se aplican a la prueba de hipótesis.

Prueba de dos colas para medias usando la distribución *t*

La especialista en recursos humanos de una importante corporación está reclutando un gran número de empleados para un proyecto en el extranjero. Durante el proceso de selección, la administración le pregunta cómo van las cosas, y ella responde: "Bien. Creo que la puntuación promedio en la prueba de aptitudes será aproximadamente 90." Cuando la administración revisa 20 de los resultados de la prueba, encuentra que la puntuación media es 84, y la desviación estándar de esta puntuación es 11.

$$\mu_{H_0} = 90 \leftarrow \text{Valores hipotéticos de la media de población}$$
$$n = 20 \leftarrow \text{Tamaño de la muestra}$$
$$\bar{x} = 84 \leftarrow \text{Media de la muestra}$$
$$s = 11 \leftarrow \text{Desviación estándar de la muestra}$$

Si la administración desea probar su hipótesis al nivel de significancia de 0.10, ¿cuál es el procedimiento a seguir?

Paso 1: Establezca sus hipótesis, tipo de prueba y nivel de significancia

$$H_0: \mu = 90 \leftarrow \text{Hipótesis nula: la puntuación media real de población es 90}$$
$$H_1: \mu \neq 90 \leftarrow \text{Hipótesis alternativa: la puntuación media no es 90}$$
$$\alpha = 0.10 \leftarrow \text{Nivel de significancia para probar esta hipótesis}$$

Paso 2: Elija la distribución apropiada y encuentre el valor crítico

La figura 8-18 ilustra la gráfica de este problema. Puesto que la administración está interesada en saber si la puntuación media verdadera es *mayor* o *menor* que la puntuación hipotética, es apropiado usar una *prueba de dos colas.* El nivel de significancia de 0.10 se indica en la figura 8-18 como las dos áreas sombreadas; cada una contiene 0.05 del área bajo la distribución *t*. Como el tamaño de muestra es 20, el número apropiado de grados de libertad es 19, es decir, $20 - 1$. Entonces, buscamos en la tabla de la distribución *t* del apéndice, en la columna de 0.10 y el renglón correspondiente a 19 grados de libertad. Encontramos que el valor crítico de *t* es 1.729.

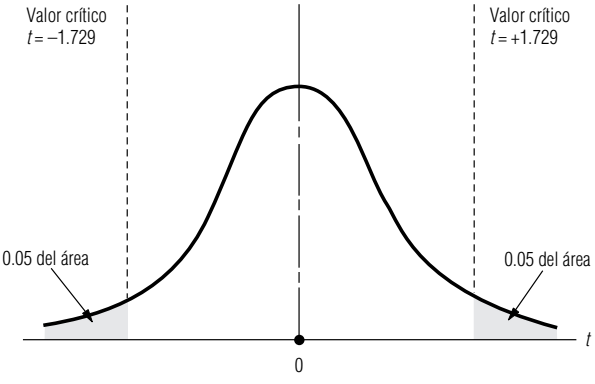


FIGURA 8-18

Prueba de hipótesis de dos colas al nivel de significancia de 0.10 usando la distribución *t*

Como no se conoce la desviación estándar de la población, debemos estimarla usando la desviación estándar de la muestra y la ecuación 7-1:

$$\hat{\sigma} = s \qquad [7\text{-}1]$$
$$= 11$$

Paso 3: Calcule el error estándar y estandarice el estadístico de la muestra

Ahora podemos calcular el error estándar de la media. Como estamos usando $\hat{\sigma}$, una estimación de la desviación estándar de la población, el error estándar de la media también será una estimación. Podemos utilizar la ecuación 7-6, como sigue:

$$\hat{\sigma}_{\bar{x}} = \frac{\hat{\sigma}}{\sqrt{n}} \qquad [7\text{-}6]$$
$$= \frac{11}{\sqrt{20}}$$
$$= \frac{11}{4.47}$$
$$= 2.46 \leftarrow \text{Error estándar estimado de la media}$$

A continuación estandarizamos la media de la muestra, $\bar{x}$, restando μ_{H_0}, la media hipotética, y dividiendo entre $\hat{\sigma}_{\bar{x}}$, el error estándar estimado de la media. Como nuestra prueba de hipótesis se basa en la distribución *t*, usamos *t* para denotar el estadístico estandarizado.

$$t = \frac{\bar{x} - \mu_{H_0}}{\hat{\sigma}_{\bar{x}}}$$
$$= \frac{84 - 90}{2.46}$$
$$= -2.44$$

Paso 4: Bosqueje la distribución y señale el valor de la muestra y los valores críticos

Paso 5: Interprete el resultado

Al señalar este resultado en una gráfica de la distribución muestral, nos damos cuenta de que la media de la muestra cae fuera de la región de aceptación, como se ve la figura 8-19.

Por tanto, la administración debe rechazar la hipótesis nula (la aseveración de la directora del departamento de personal acerca de que la puntuación media real de los empleados examinados es 90).

Pruebas de una cola para medias usando la distribución *t*

Una diferencia respecto a las tablas *z*

El procedimiento para una prueba de hipótesis de una cola utilizando la distribución *t* es el mismo, conceptualmente, que para una prueba de una cola usando la distribución normal y la tabla *z*. Sin embargo, realizar esas pruebas de una cola puede conllevar algunas dificultades. Observe que los enca-

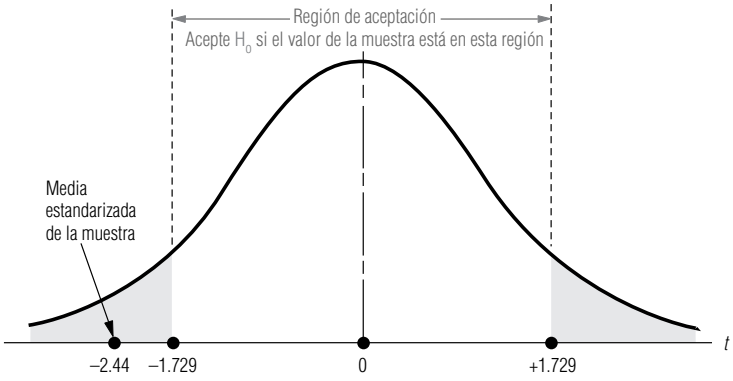


FIGURA 8-19

Prueba de hipótesis de dos colas al nivel de significancia de 0.10; se observa la región de aceptación y la media estandarizada de la muestra

bezados de las columnas de la tabla 2 del apéndice representan el *área en ambas colas combinada.* Por tanto, es apropiado usarlos en una prueba de dos colas con *dos* regiones de rechazo.

Uso de la tabla *t* para pruebas de una cola

Si utilizamos la distribución *t* para una prueba de una cola, necesitamos determinar el área localizada sólo en una de ellas. Así, para encontrar el valor *t* apropiado para una prueba de una cola a un nivel de significancia de 0.05, con 12 grados de libertad, buscaríamos en la tabla 2 del apéndice en la columna con 0.10 el renglón de 12 grados de libertad. La respuesta en este caso es 1.782. **Esto es cierto porque la columna 0.10 representa 0.10 del área bajo la curva contenida en *ambas colas combinadas,* y por ende, también representa 0.05 del área bajo la curva contenida en cada cola por separado.**

Lo que viene

En el siguiente capítulo continuaremos nuestro trabajo sobre pruebas de hipótesis, con el estudio de situaciones en las que deben tomarse decisiones con base en dos muestras que pueden provenir o no de la misma población.

SUGERENCIAS Y SUPOSICIONES

Realizar pruebas de hipótesis con la distribución *t* no es diferente a hacerlo con la distribución normal, excepto porque se usa una tabla distinta y debe darse el número de grados de libertad. Sugerencia: el número de grados de libertad en una prueba de una sola muestra es siempre una unidad menos que el tamaño de la muestra. Advertencia: use la distribución *t* siempre que el tamaño de la muestra sea menor que 30, la desviación estándar de la población no se conozca y la población sea normal o aproximadamente normal.

Ejercicios 8.7

Ejercicios de autoevaluación

EA 8-11 Dado que la media de la muestra es 83, la desviación estándar de la muestra es 12.5 y el tamaño de la muestra es 22, pruebe la hipótesis de que el valor de la media de la población es 70 contra la hipótesis alternativa de que es mayor que 70. Use un nivel de significancia de 0.025.

EA 8-12 Picosoft, Ltd., un proveedor de sistemas operativos para computadoras personales, planea la oferta pública inicial de sus acciones a fin de reunir suficiente capital de trabajo para financiar el desarrollo de un sistema integrado de la séptima generación radicalmente nuevo. Con los ingresos actuales de $1.61 por acción, Picosoft y sus aseguradores contemplan un precio de oferta de $21, o cerca de 13 veces los ingresos. Para verificar si este precio es apropiado eligieron al azar siete empresas de software en el mercado de valores y encontraron que su cociente promedio precio/utilidades era 11.6 y la desviación estándar de la muestra era 1.3. Para $\alpha = 0.02$, ¿puede Picosoft concluir que las acciones de las empresas de software en el mercado de valores tiene una razón promedio P/U que es significativamente diferente de 13?

Conceptos básicos

8-44 Dada una media de la muestra de 94.3, una desviación estándar de la muestra de 8.4 y un tamaño de muestra de 6, pruebe la hipótesis de que el valor de la media de la población es 100, contra la hipótesis alternativa de que es menor que 100. Use un nivel de significancia de 0.05.

8-45 Si una muestra de 25 observaciones revela una media muestral de 52 y una varianza muestral de 4.2, pruebe la hipótesis de que la media de la población es 65, contra la hipótesis alternativa de que es algún otro valor. Use el nivel de significancia de 0.01.

Aplicaciones

8-46 La corredora de bienes raíces Elaine Snyderman tomó una muestra aleatoria de 12 hogares de un prestigiado suburbio de Chicago y encontró que el valor de mercado promedio estimado era $780,000, con una desviación estándar de $49,000. Pruebe la hipótesis de que para todas las casas del área, el valor estimado medio es $825,000, hipótesis alternativa de que es menor que $825,000. Utilice el nivel de significancia de 0.05.

■ **8-47** Para una muestra de 60 mujeres, tomadas de un población de más de 5,000 inscritas en un programa de reducción de peso de una cadena nacional de balnearios de aguas termales, la presión sanguínea diastólica media de la muestra es 101 y la desviación estándar de la muestra es 42. Para un nivel de significancia de 0.02, ¿puede concluir que, en promedio, las mujeres inscritas en el programa tienen una presión sanguínea diastólica que excede el valor de 75 recomendado por diversas sociedades médicas?

■ **8-48** El departamento de procesamiento de datos de una compañía de seguros grande instaló nuevas terminales de video de color para reemplazar las unidades monocromáticas que tenían. Los 95 operadores capacitados para usar las nuevas máquinas promediaron 7.2 horas antes de lograr un desempeño satisfactorio. Su varianza muestral fue 16.2 horas al cuadrado. La larga experiencia de los operadores con las viejas terminales monocromáticas indicaba un promedio de 8.1 horas en las máquinas antes de que su desempeño fuera satisfactorio. Al nivel de significancia de 0.01, ¿debería el supervisor del departamento concluir que es más fácil aprender a operar las nuevas terminales?

■ **8-49** Con la caída del mercado petrolero de principios de 1986, los educadores de Texas se preocuparon por la forma en que las pérdidas resultantes en los ingresos del Estado (estimadas en cerca de $100 millones por cada disminución de un dólar en el precio del barril de petróleo) afectarían sus presupuestos. La directiva estatal de educación pensaba que la situación no sería crítica en tanto pudieran estar razonablemente seguros de que el precio permanecería arriba de $18 por barril. Encuestaron a 13 economistas especializados en el mercado del petróleo, elegidos al azar, y les pidieron que predijeran qué tanto bajarían los precios antes de repuntar. Las 13 predicciones promediaron $21.60, con una desviación estándar de $4.65. Para un nivel $\alpha = 0.01$, ¿es la predicción promedio significativamente mayor que $18.00? ¿Debe la directiva de educación concluir que es improbable una crisis presupuestaria? Explique su respuesta.

■ **8-50** Un documental televisivo acerca de comer en exceso afirmaba que los estadounidenses tienen un sobrepeso aproximado de 10 libras en promedio. Para probar esta afirmación, examinaron a 18 individuos elegidos aleatoriamente, y encontraron que su sobrepeso promedio era 12.4 libras, con una desviación estándar de la muestra de 2.7 libras. A un nivel de significancia de 0.01, ¿hay alguna razón para dudar de la validez del valor afirmado de 10 libras?

■ **8-51** XCO, un fabricante multinacional, usa un proceso por lotes para producir objetos. La producción de cada lote toma 8 horas y tiene costos de material y mano de obra de $8,476. Debido a la variación en la eficiencia de la máquina y la pureza del material, el número de objetos por lote es aleatorio. Todos los objetos se pueden vender en $2.50 cada uno y la producción es rentable siempre que los lotes se vendan en más de $12,500 en promedio. XCO hizo un muestreo de 16 lotes y encontró 5,040 objetos por lote en promedio, con una desviación estándar de 41.3 objetos. Para $\alpha = 0.025$, ¿puede XCO concluir que su operación de objetos es rentable?

Soluciones a los ejercicios de autoevaluación

EA 8-11 $s = 12.5 \quad n = 22 \quad \bar{x} = 83$

$H_0: \mu = 70 \quad H_1: \mu > 70 \quad \alpha = 0.025$

El límite superior de la región de aceptación es $t = 2.080$, o

$$\bar{x} = \mu_{H_0} + ts/\sqrt{n} = 70 + 2.080(12.5)/\sqrt{22} = 75.54$$

Debido a que el valor t observado $= \dfrac{\bar{x} - \mu_{H_0}}{s/\sqrt{n}} = \dfrac{83 - 70}{12.5/\sqrt{22}} = 4.878 > 2.080$ (o $\bar{x} > 75.54$), se rechaza H_0.

EA 8-12 $s = 1.3 \quad n = 7 \quad \bar{x} = 11.6$

$H_0: \mu = 13 \quad H_1: \mu \neq 13 \quad \alpha = 0.02$

El límite superior de la región de aceptación es $t = \pm 3.143$, o

$$\bar{x} = \mu_{H_0} + ts/\sqrt{n} = 13 \pm 3.143(1.3)/\sqrt{7} = (11.46, 14.54)$$

Debido a que el valor t observado $= = \dfrac{\bar{x} - \mu_{H_0}}{s/\sqrt{n}} = \dfrac{11.6 - 13}{1.3/\sqrt{7}} = -2.849 > -3.143$ (o $\bar{x} = 11.6$, que esta entre 11.46 y 14.54), no se rechaza H_0. El cociente promedio P/U de las empresas de software en el mercado de valores no es significativamente diferente a 13.

Estadística en el trabajo

Loveland Computers

Caso 8: Pruebas de hipótesis de una muestra "Ésta es la otra cosa que me ha puesto a pensar más sobre la posibilidad de abrir un departamento de software", dijo Margot Derby, la jefa de comercialización de Loveland Computers, mientras sacaba una columna de *The Wall Street Journal* del cajón de su escritorio. "Como sabes, los precios de las PC han estado bajando. Pero, para sorpresa de todos, los compradores parecen estar gastando lo mismo en total; están compensando el precio de descuento con la compra de más *monerías* y más software.

"El artículo cita una cifra de la cantidad promedio gastada en software durante el primer año después de la adquisición de una máquina. Ésa es la misma cifra que buscábamos cuando hicimos nuestra encuesta telefónica, pero nuestro número salió mucho menor que la cantidad en el artículo. El problema es que no estoy segura de cuál cifra usar para hacer el plan comercial para un nuevo departamento de software."

"Bueno, ¿por qué saldría tu número diferente?", preguntó Lee Azko.

"No pretendemos atraer a todo el mundo", respondió Margot. "Tal vez tenemos una imagen más 'técnica', y por consiguiente, nuestros clientes pueden ser diferentes al cliente 'promedio' del que hablan en ese artículo. Tal vez utilizan programas hechos por ellos mismos."

"O tal vez la diferencia no significa nada y es sólo el resultado de un error de muestreo", sugirió Lee.

"Pero no sé cómo decidir con seguridad. Hemos calculado la media y la desviación estándar de nuestra muestra telefónica, pero el artículo del *Journal* sólo nos da la media. Y recuerdo lo suficiente de mi curso de estadística de la universidad para saber que no podemos hacer una prueba si no conocemos la desviación estándar de la población.

Preguntas de estudio: Suponga que la cifra media de gastos de software citada en el periódico es una *media de población* confiable. ¿Tiene razón Margot en que Lee también debe conocer la desviación estándar de la población para realizar una prueba? ¿Qué idea es la que Margot está explorando aquí? ¿Cómo se establecería esta idea en términos de pruebas de hipótesis?

Ejercicio de base de datos computacional

HH Industries

Hal pasó a ver a Laurel al día siguiente de la reunión del personal. "Tengo una pregunta", empezó diciendo, "es sobre el estudio que hicieron tú y Bob respecto a la licitación competitiva de las órdenes de compra. Sé que soy un novato en estadística y estoy esforzándome por entender los intervalos de confianza y esas cosas, pero, ¿no hay alguna manera de obtener un sí o un no a la pregunta que hicimos?", hizo una pausa y prosiguió: "no quise hacer preguntas tontas en la reunión de ayer, porque realmente siento como si mi falta de conocimientos fuera lo que me hace cuestionar los resultados. Trata de verlo desde mi punto de vista: la cifra que quiero, 60%, está *en* el intervalo de confianza, pero enfrentémoslo, la mayor parte del intervalo es *menor* que 0.6. ¿Ves a lo que me refiero?"

Laurel asintió y sonrió. "Entiendo exactamente lo que quieres decir", dijo. "No eres tan novato como piensas y no es una pregunta tonta. Los intervalos de confianza son, por naturaleza, algo confusos. Tengo un libro de texto que me gustaría prestarte para que tengas un poco más de información. Y en cuanto a la respuesta de sí o no, déjame que te presente a las pruebas de hipótesis." Laurel prosiguió explicándole las pruebas de hipótesis de una y dos colas, ilustrándolas con la situación que tenían a la mano.

1. Efectúe una prueba de hipótesis de una cola para determinar si la proporción de órdenes de compra licitadas en forma competitiva es realmente menor que 0.6, usando los valores $\overline{p} = 0.58$ y $n = 1{,}052$. Pruebe al nivel de significancia de 0.01.

Gary le abrió la puerta a Laurel una mañana. "No te he visto en un buen tiempo, exceptuando nuestras reuniones semanales de personal", dijo sonriendo. "De hecho, creo que fue desde que me ayudaste con el estudio de códigos postales, poco después de que llegaste aquí. ¡He oído que te tienen ocupada!"

"Ha estado bastante agitado", convino Laurel. "Pero la estoy pasando bien. ¡Estar ocupada es mejor que estar aburrida! Aunque hace poco, el ritmo bajó un poco, lo cual está bien, y ya pronto serán mis vacaciones. Voy a esquiar en las Montañas Rocallosas, lo que constituirá un buen descanso."

"¡Suena divertido! ¿Crees tener tiempo antes de irte para ayudarme con un pequeño problema?", Laurel asintió. Gary continuó, "hemos estado observando algunos indicadores de un mal control de calidad de uno de nuestros proveedores de sellos O-Ring. El ajuste de este tipo particular de sello es crucial, como tal vez sepas. En consecuencia, cuando Stan me dice que uno o dos de nuestros clientes se han quejado, necesitamos hacer una verificación cuidadosa".

"¿Tienen procedimientos de inspección al recibir?", preguntó Laurel, al tiempo que entraba a su oficina.

"Buena pregunta", Gary hizo una mueca. "Hacemos lo que podemos, e inspeccionamos cada una de las refacciones grandes en pequeñas cantidades. Pero cosas como los sellos,

nos llegan miles al mismo tiempo y a menudo todo lo que podemos hacer es inspeccionar un puñado. Realmente me gustaría establecer procedimientos más rigurosos como un objetivo a largo plazo."

"Parece que el problema de los sellos O-Ring es un buen lugar para empezar. Te informaré de los datos que vamos a necesitar y los pasos del análisis, y tal vez puedas partir de allí para evaluar otros productos." Laurel, sabía que poco a poco se ganaba la confianza del personal, no sólo en cuanto a su capacidad, sino en especial con respecto a la validez y a la utilidad del análisis estadístico. "Me voy a dar una vuelta por tu oficina al rato para obtener más detalles."

Más tarde, Laurel encontró a Gary en el taller. "Eres muy oportuna", dijo Gary. "Éste es el último lote, está llegando en este momento. ¿Ves el tamaño?", sacó uno de los sellos de la caja y señaló la abertura interna. "El promedio requerido es 0.140 pulg. Históricamente hemos tenido una desviación estándar de 0.003 pulg. Tenemos un instrumento de calibración para medirlos. ¿Y si les pido a mis muchachos que registren los datos de este lote y que te los entreguen en tu oficina mañana?"

"Eso estaría bien", respondió Laurel. "Sólo para mi propia información, ¿es igual de malo que queden demasiado justos o demasiado sueltos?"

"Sip. De cualquier forma puede resultar desastroso para el usuario", le explicó Gary.

2. A un nivel de significancia de 0.05, ¿debería concluir Laurel que los sellos O-Ring están fuera de tolerancia? Use los datos proporcionados en los archivos CH08.xxx del CD que acompaña al libro.

Del libro de texto al mundo real

Prueba de hipótesis: impacto de la relajación del reglamento en la administración de ferrocarriles

Los momentos de cambio a menudo estimulan los estudios estadísticos. Éste fue el caso de la industria ferrocarrilera durante la década de los años 80 en Estados Unidos, cuando el Decreto de Eficiencia y la desregulación que le siguió incrementaron la competencia en una industria antes protegida y estancada. En 1987, los analistas probaron la hipótesis general de que las características observables y la estructura de la administración de ferrocarriles habían comenzado a cambiar junto con la reciente reforma al reglamento. Se analizaron las características de administración para 1977 y 1983. El estudio indicó que los administradores eran significativamente más jóvenes y mejor preparados, con menos años de servicio en compañías e industrias. Estos resultados mostraron que la desregulación tuvo éxito en rejuvenecer una industria decadente, enfrentándola al reto de atraer administradores que pudieran adaptarse a un entorno cambiante.

Antecedentes Desde principios del siglo XX, los niveles de empleo y el porcentaje de tráfico de carga dentro de Estados Unidos han caído drásticamente. Debido a que la reglamentación federal altamente restrictiva permitía que la estructura de administración de los ferrocarriles estadounidenses fuera estática y burocrática, la industria tuvo bajos niveles de rentabilidad y competitividad. Históricamente, pocos administradores de ferrocarriles provenían de fuera de la industria. Para mejorar los prospectos financieros se relajó el reglamento de la industria durante los ochenta.

Prueba de hipótesis Curtis Grimm, James Kling y Ken Smith de la Universidad de Maryland probaron la hipótesis de que la desregulación forzaría una modernización de las actitudes administrativas. Se probaron varias hipótesis específicas para determinar la validez de las aseveraciones gubernamentales de que la desregulación mejoraba la competitividad. Algunas de las hipótesis probadas son:

I. Se supuso que la edad promedio de los administradores en 1983 era significativamente menor que en 1977. Un entorno competitivo requiere que los ferrocarriles recluten administradores más jóvenes con mayor probabilidad de ser más innovadores y agresivos. Se dijo que la edad era un indicador de la flexibilidad propia y de la propensión a aceptar riesgos.

2. La segunda hipótesis referente a los administradores era que el número promedio de años de empleo con su actual empleador ferrocarrilero debería ser significativamente menor en 1983 que en 1977. Los años de experiencia pueden actuar como un obstáculo para el cambio organizacional en una industria con nueva reglamentación.

3. También se supuso que el administrador promedio tendría menos años de empleo en la industria ferrocarrilera en 1983 que en 1977. La desregulación requeriría que los ferrocarriles buscaran administradores con experiencia en entornos competitivos.

Resultados Los resultados del estudio de Maryland demostraron que la administración de ferrocarriles había cambiado como respuesta a la reforma al reglamento. La edad promedio de los administradores de ferrocarriles declinó en más de 1 año completo, de 51.68 años en 1977 a 50.66 años en 1983. Esta tendencia hacia una administración más joven respaldó la primera hipótesis. Se encontró que el ejercicio promedio de los administradores de ferrocarriles era 25 años en 1977 y 21.33 años en 1983. Esta caída de 3.77 años era significativa aun con α tan pequeña como 0.001 y respaldó la segunda hi-

pótesis. La caída comparativamente mayor en el tiempo de servicio contra la edad indicó que sólo la llegada de administradores más jóvenes no era responsable de la disminución de los años de servicio. Los ferrocarriles evidentemente ahora incluían gerentes de fuera de la compañía con más frecuencia que en el pasado. También se reveló un fuerte respaldo para la tercera hipótesis. El número promedio de años en la industria declinó de 27.23 años en 1977 a 24.88 años en 1983, estadísticamente significativo al nivel de 0.001. Es claro que, los resultados respaldaron la hipótesis de que la administración de ferrocarriles sería más joven y cambiante después de la desregulación.

Conclusiones Los ferrocarriles, las aerolíneas y las compañías de transportes motorizados han experimentado desregulaciones sustanciales durante la última década, mismas que crearon un ímpetu por el cambio administrativo dentro de estas industrias. La prueba de hipótesis y otros métodos estadísticos proporcionan un medio para determinar los efectos de la reforma regulatoria en las industrias. Esta investigación puede ser benéfica para el desarrollo de estructuras administrativas exitosas en las compañías de todas las industrias, en la búsqueda de prosperidad en entornos altamente competitivos.

Fuente: Curtis M. Grimm, James A. Kling y Ken G. Smith, "The Impact of U.S. Rail Regulatory Reform on Railroad Management and Organizational Structure", en *Transportation Research-A* 21A(2), (1987): 87-94.

Repaso del capítulo

● Términos introducidos en el capítulo 8

Alfa (α) Probabilidad de cometer un error tipo I.

Beta (β) Probabilidad de cometer un error tipo II.

Curva de potencia Gráfica de los valores de la potencia de una prueba para cada valor de μ, u otro parámetro de población para el que la hipótesis alternativa es cierta.

Error tipo I Rechazo de una hipótesis nula cuando es verdadera.

Error tipo II Aceptación de una hipótesis nula cuando es falsa.

Escala estandarizada Medición expresada en desviaciones estándar a partir de la media de la variable.

Escala sin procesar Medición en las unidades originales de la variable.

Hipótesis Suposición o especulación que hacemos respecto a un parámetro de población.

Hipótesis alternativa Conclusión que aceptamos cuando los datos no respaldan la hipótesis nula.

Hipótesis nula Hipótesis o suposición respecto a un parámetro de población que deseamos probar, generalmente una suposición sobre el *status quo*.

Nivel de significancia Valor que indica el porcentaje de valores de la muestra que están fuera de ciertos límites, suponiendo que la hipótesis nula es correcta; es decir, la probabilidad de rechazar la hipótesis nula cuando es cierta.

Potencia de la prueba de hipótesis Probabilidad de rechazar la hipótesis nula cuando es falsa, es decir, una medida de qué tan bien funciona la prueba de hipótesis.

Prueba de dos colas Prueba de hipótesis en la que la hipótesis nula se rechaza si el valor de la muestra es significativamente mayor o menor que el valor hipotético del parámetro de población; prueba que involucra dos regiones de rechazo.

Prueba de cola derecha (o prueba de cola superior) Prueba de hipótesis de una cola en la que un valor de la muestra significativamente mayor que el valor hipotético de la población nos llevará a rechazar la hipótesis nula.

Prueba de cola izquierda (o prueba de cola inferior) Prueba de hipótesis de una cola en la que un valor de la muestra que es significativamente menor que el valor hipotético de la población nos llevará a rechazar la hipótesis nula.

Prueba de una cola Prueba de hipótesis en la que sólo hay una región de rechazo, es decir, sólo nos interesa si el valor observado se desvía del valor hipotético en una dirección.

Valor crítico Valor del estadístico estándar (z o t) más allá del cual rechazamos la hipótesis nula; el límite entre las regiones de aceptación y de rechazo.

● Ejercicios de repaso

■ **8-52** Para las situaciones siguientes establezca las hipótesis nula y alternativa apropiadas.

a) La Oficina del Censo de Estados Unidos desea determinar si el porcentaje de personas que carecen de hogar en la ciudad de Nueva York es el mismo que el promedio nacional.

b) El dueño de una ferretería local desea determinar si las ventas de herramientas para jardín resultaron mejores de lo normal después de una promoción de primavera.

c) El canal del clima desea saber si el promedio de nevadas en los años 80 fue significativamente diferente del promedio de 8 pulg registrado en los últimos 100 años.

d) Una revista para el consumidor se cuestiona acerca de si el rendimiento de combustible de un nuevo automóvil subcompacto es significativamente menor que las 34 millas por galón que anuncia la etiqueta que lleva pegada a la ventanilla.

■ **8-53** La compañía Health Electronics, Inc., fabricante de baterías para marcapasos, especifica que la vida de cada batería es mayor o igual que 28 meses. Si el calendario de cirugía para remplazar las baterías se ha de basar en esta afirmación, explique a la administración de esta empresa las consecuencias que se tendrían al cometer errores tipo I y tipo II.

■ **8-54** Un fabricante de ropa deportiva para mujeres pequeñas ha supuesto que el peso promedio de las mujeres que compran su ropa es 110 libras. La compañía tomó dos muestras de sus clientas y encontró que una estimación muestral de la media de la población es 98 libras, y que la otra muestra produce un peso medio de 122 libras. En la prueba de hipótesis de la compañía de que la media de la población es 110 libras contra la hipótesis de que la media no es igual a 110 libras, ¿es más probable que alguno de estos valores de la muestra nos lleve a aceptar la hipótesis nula? ¿Por qué sí o por qué no?

■ **8-55** Muchas ciudades estadounidenses han establecido carriles para Vehículos de Alta Ocupación (VAO) para acelerar el transporte hacia los distritos de negocios del centro. Para la autoridad de tráfico, la planeación de tránsito ha dependido de un bien establecido promedio de 3.4 ocupantes por VAO. Sin embargo, un becario de la institución observa que dado que muchas empresas están patrocinando transporte corporativo en camionetas, el número promedio de pasajeros por vehículo puede ser mayor. El becario tomó una muestra de 23 vehículos que pasan por el carril VAO en una caseta de pago y reporta una media muestral de 4.3 pasajeros con una desviación estándar de 1.5 pasajeros. Al nivel 0.01 de significancia, ¿sugiere la muestra que la media del número de pasajeros se ha incrementado?

■ **8-56** En el ejercicio EV 8-5, ¿cuál sería la potencia de la prueba para $\mu = 14{,}000$, 13,500 y 13,000 si el nivel de significancia se cambia a 0.10?

■ **8-57** En un día promedio, alrededor del 5% de las acciones de la Bolsa de Valores de Nueva York muestran una nueva alza para ese año. El viernes 18 de septiembre de 1992, el promedio industrial Dow Jones cerró en 3,282 con un fuerte volumen de más de 136 millones de títulos negociados. Una muestra aleatoria de 120 títulos determinó que 16 de ellos habían mostrado nuevas alzas anuales ese día. Usando un nivel de significancia de 0.01, ¿deberíamos concluir que más títulos de los habituales tuvieron nuevas alzas anuales ese día?

■ **8-58** En respuesta a las críticas por correspondencia perdida, el Servicio Postal estadounidense puso en marcha nuevos procedimientos para disminuir el problema. Al director general de correos se le aseguró que dicho cambio reduciría los extravíos a menos del histórico índice de pérdidas del 0.3%. Después de dos meses de implantar los nuevos procedimientos, el servicio postal patrocinó una investigación en la que se envió un total de 8,000 piezas de correspondencia desde diferentes partes del país. Dieciocho de estas piezas de prueba no llegaron a su destino. A un nivel de significancia de 0.10, ¿puede el director general de correos concluir que los nuevos procedimientos consiguieron su objetivo?

■ **8-59** ¿Cuál es la probabilidad de que estemos rechazando una hipótesis nula verdadera cuando rechazamos el valor hipotético debido a que

a) La estadística de muestra difiere del valor hipotético en más de 2.15 errores estándar en cualquier dirección.

b) El valor del estadístico de la muestra es mayor en más de 1.6 errores estándar.

c) El valor del estadístico de la muestra es menor que el valor hipotético en más de 2.33 errores estándar.

Los distribuidores de fondos de inversión mutua *abiertos* venden acciones adicionales a los individuos que desean invertir en estos fondos. Estas acciones están valuadas como *valores de activo neto*, el valor de estos títulos representado por las acciones de los fondos, más una comisión (o *cargo*) en el intervalo de 0 a 8%. Los fondos *cerrados*, por otra parte, tienen un número fijo de acciones. Estas acciones se negocian en varias bolsas de valores a precios determinados por el mercado. Si el precio es mayor que el valor activo neto, se dice que la acción se vende con *prima*: si sucede lo contrario se vende con *descuento*. Una muestra aleatoria de 15 fondos cerrados el 4 de junio de 1993, encontró los descuentos (valores negativos) y las primas (valores positivos) expresados en porcentajes y enumerados en la tabla MR8-1. Utilice esta información para responder los ejercicios 8-60 y 8-61.

■ **8-60** ¿Se venden significativamente menos fondos de acciones de interés variable con descuentos que con primas? Pruebe a un nivel $\alpha = 0.01$.

Tabla MR8-1
Descuentos y primas para una muestra de 15 fondos mutualistas cerrados

Nombre del título	Descuento/prima
Blue Chip Value	+4.7
Gabelli Equity Trust	−0.7
Liberty AII-Star	+5.3
Central Fund of Canada	+9.2
Global Health Sciences	−0.3
Patriot Global Dividend	−0.3
Preferred Income	+5.0
Austria Fund	+0.4
Emerging Mexico	−1.9
First Australia	+0.5
Germany Fund	+5.8
Japan Equity	−8.2
Latin America Equity	−9.4
Morgan Stanley Emerging Markets	+10.3
Turkish Investment	+1.7

Fuente: The Wall Street Journal *(17 de junio de 1993): B8.*

■ **8-61** Un profesor de finanzas desarrolló una teoría que predice que los fondos de acciones de interés variable cerrados se deberían vender con una prima cercana al 5% en promedio. suponiendo que la población descuento/prima tiene una distribución aproximadamente normal, ¿apoya la información muestreada esta teoría? Pruebe con $\alpha = 0.05$.

■ **8-62** ¿Si deseamos aceptar la hipótesis nula como correcta el 85% de las veces, ¿a cuánto errores estándar alrededor de la media hipotética debe encontrarse la media muestral para estar en la región de aceptación? ¿Y si deseamos estar 98% seguros de aceptar la hipótesis nula cuando es cierta?

■ **8-63** Los estatutos ambientales federales aplicables a cierta planta nuclear especifican que el agua reciclada no debe, en promedio, exceder los 84°F (28.9°C) antes de que se arroje al río que corre junto a la planta. De 70 muestras, se encontró que la temperatura promedio del agua reciclada era 86.3°F (30.2°C) Si la desviación estándar de la población es 13.5°F (7.5°C), ¿debería la planta ser multada por exceder las limitaciones del estatuto? Formule y pruebe las hipótesis apropiadas con $\alpha = 0.05$.

■ **8-64** Inspectores del gobierno, al investigar los cargos levantados contra una embotelladora de refrescos de Lousiana, que no llenaba bien sus productos, muestrearon 200 botellas y encontraron que el promedio de llenado es 31.7 onzas líquidas. Se anuncia que las botellas contienen 32 onzas líquidas. Se sabe que la desviación estándar de le población es 1.5 onzas líquidas. ¿Deberían concluir los inspectores, a un nivel de significancia del 2%, que las botellas están tienen menos contenido?

■ **8-65** En 1995, la tarifa aérea promedio con dos semanas de anticipación en el vuelo entre Raleigh-Durham, Carolina del Norte y la ciudad de Nueva York era $235. La desviación estándar de la población era $68. En una encuesta hecha en 1996 a 90 viajeros, elegidos al azar, entre estas dos ciudades se encontró que habían pagado en promedio, $218.77 por sus boletos. ¿Cambió significativamente la tarifa aérea promedio en esta ruta entre 1995 y 1996? ¿Cuál es el mayor valor de alfa para el que podría concluir que la tarifa promedio observada no es significativamente diferente de $235?

■ **8-66** Audio Sounds maneja una cadena de tiendas que venden sistemas y componentes estéreo de audio. Ha tenido mucho éxito en muchos pueblos universitarios, pero también algunos fracasos. El análisis de estas fallas la ha llevado a adoptar la política de no abrir una tienda a menos que estén razonablemente seguros de que al menos el 15% de los estudiantes del lugar tienen sistemas estéreo con un costo de $1,100 o más. En una encuesta hecha a 300 de los 2,400 estudiantes de una pequeña escuela de artes en el medio oeste de Estados Unidos se descubrió que 57 de ellos tienen un sistema estéreo con un costo mayor o igual que $1,100. Si Audio Sounds desea correr un riesgo de fracaso del 5%, ¿debería abrir una tienda en este lugar?

■ **8-67** La ciudad de Oakley recoge un impuesto de transferencia del 1.5% al cerrar las transacciones de bienes raíces. En una semana promedio, suelen cerrarse 32 transacciones, con una desviación estándar de 2.4. Con un nivel de significancia de 0.10, ¿estaría de acuerdo con la conclusión del recaudador de impuestos de que “las ventas están fuera de lugar este año” si una muestra de 16 semanas tiene una media de 28.25 transacciones cerradas?

■ **8-68** En 1996, se estimó que alrededor del 72% de los hogares de Estados Unidos es suscriptor de televisión por cable. Los editores de la revista *Newstime* estaban seguros de que sus lectores tenían suscripción por

cable en un promedio más alto que la población en general y querían usar este hecho para ayudar a vender espacio de publicidad para los canales de estreno en cable. Para verificar esto, muestrearon a 250 de los suscriptores del *Newstime* y encontraron que 194 de ellos tenían suscripción a televisión por cable. Con un nivel de significancia del 2%, ¿los datos de la encuesta apoyan el parecer de los editores?

8-69 Una compañía, recientemente criticada por no pagar lo mismo a hombres que a mujeres que realizan el mismo trabajo, declara que el sueldo promedio pagado a todos los empleados es $23,500 anuales. De una muestra aleatoria de 29 mujeres que laboran en la compañía, se calculó que el salario promedio era $23,000. Si se sabe que la desviación estándar de la población es $1,250 para estos trabajos, determine si es razonable esperar (con dos errores estándar) que la media de la muestra sea $23,000 si, en efecto, es cierto lo declarado por la compañía.

8-70 Drive-a-Lemon, renta automóviles en buenas condiciones mecánicas, pero más antiguos que los que ofrecen las grandes cadenas nacionales de renta de coches. Como resultado, anuncia que sus tarifas son considerablemente más bajas que las de sus competidores grandes. Una encuesta de la industria estableció que el cargo total promedio por renta en una de las compañías más importantes es $77.38. Una muestra aleatoria de 18 transacciones realizadas por Drive-a-Lemon mostró un cargo total promedio de $87.61, con una desviación estándar muestral de $19.48. Verifique que para $\alpha = 0.025$, el cargo total promedio de Drive-a-Lemon es significativamente *más alto* que el de las compañías grandes. ¿Indica este resultado que las tarifas de Drive-a-Lemon, en efecto, no son más bajas que las de las cadenas nacionales importantes? Explique su respuesta.

Una muestra aleatoria de 20 corporaciones privadas de Carolina del Norte reveló los datos de la tabla MR8-2 acerca de sus directores ejecutivos. Use esta información para responder a los ejercicios 8-71 a 8-73.

8-71 Suponiendo que la población tiene una distribución aproximadamente normal, ¿es razonable concluir, con un nivel de significancia del 5%, que la edad promedio de los directores ejecutivos de todas las corporaciones de Carolina del Norte es significativamente mayor que 50 años?

8-72 La mediana del número de niños en los hogares de Estados Unidos es de 2. Con un nivel de significancia del 10%, ¿debería concluir que la proporción de familias de directores ejecutivos con más de 2 niños es significativamente diferente de 0.5?

8-73 Alrededor del 65% de todos los adultos de Estados Unidos son casados. ¿La proporción de directores ejecutivos casados es significativamente mayor que la de la población general? Pruebe con $\alpha = 0.02$.

8-74 Regrese al ejercicio 8-26 y calcule la potencia de la prueba para $\mu =$ $41.95, $42.95 y $43.95.

8-75 Un jefe de personal piensa que el 18% de los empleados de la compañía trabaja horas extra cada semana. Si la proporción observada esta semana es 13% en una muestra de 250 de los 2,500 empleados, ¿podemos aceptar que su opinión es razonable o debemos concluir que es más apropiado algún otro valor? Use $\alpha = 0.05$.

8-76 Regrese al ejercicio EV 8-5 y calcule la potencia de la prueba para $\mu =$ 14,000, 13,500 y 13,000.

8-77 Una corredora de bolsa afirma que puede predecir, con el 85% de certeza, el ascenso o caída, durante el mes siguiente, de un valor del mercado de valores. Para probarlo, predice el resultado de 60 valores y acierta en 45 de sus predicciones ¿Presentan estos datos evidencia concluyente (con $\alpha = 0.04$) de que la exactitud de sus predicciones es significativamente menor que el 85% declarado?

8-78 En el ejercicio 8-26, ¿cuál sería la potencia de la prueba para $\mu =$ $41.95, $42.95 y $43.95, si el nivel de significancia se cambiara a 0.05?

8-79 Un fabricante de suplementos vitamínicos para recién nacidos incluye un cupón para una muestra gratis de este producto en la canasta distribuida a los nuevos padres en los hospitales. Históricamente, cerca del 18% de los cupones se canjean. Dada la tendencia actual de tener menos hijos e iniciar una familia más tarde, la empresa se imagina que los padres actuales están mejor informados en promedio y, como resultado, es más probable que utilicen un suplemento vitamínico para sus hijos. Una muestra de 1,500 nuevos padres canjeó 295 cupones. ¿Respalda este hecho, a un nivel de significancia del 2%, la opinión de la compañía respecto a los nuevos padres actuales?

8-80 Un innovador de la industria automotriz pensó que su nuevo automotor eléctrico captaría el 48% del mercado regional en un año, debido al bajo precio y el rendimiento superior de su producto. Existen 5,000 usuarios de automotores en la región. Después de muestrear el 10% de estos usuarios un año después de lanzar su producto al mercado, la compañía encontró que el 43% de ellos usaba los nuevos motores. Con un nivel $\alpha = 0.01$, ¿deberíamos concluir que la compañía fracasó en alcanzar su objetivo de participación en el mercado?

Tabla MR8-2
Datos personales de una muestra de 35 directores ejecutivos

Nombre de la compañía	Edad	Estado	Niños
Parkdale Milis Inc.	68	M	3
SAS Institute Inc.	50	M	3
Cogentrix Inc.	65	M	3
House of Raeford Farms Inc.	66	M	3
Harriet & Henderson Yarns Inc.	52	M	1
Harvey Enterprises and Affiliates	44	M	4
Radiator Specialty Co.	77	M	3
Parrish Tire Co.	43	M	2
Spectrum Dyed Yarns Inc.	59	M	2
Southeastern Hospital Supply Corp.	45	M	4
Miller Building Corp.	55	M	3
Pneumafil Corp.	55	S	O
Kroehler Furniture Industries Inc.	50	M	3
Carolina Petroleum Distributors Inc.	42	D	2
Tanner Cos.	64	M	4
Raycom Inc.	43	M	2
Cummins Atlantic Inc.	57	M	4
W. R. Bonsal Co.	62	M	3
Maola Milk & Ice Cream Co.	67	M	2
Waste Industries Inc.	56	M	2

Estado = estado civil (soltero, **S**; casado, **M**; o divorciado, **D**)
Niños = número de hijos

Fuente: "Geffing a Grip an Clasely Held Campanies", Bussines North Carolina *13(2) (junio de 1993): 28-63.*

■ **8-81** Según las especificaciones, las máquinas tragamonedas de palanca en los casinos deben pagar una vez cada 11.6 giros, con una desviación estándar de 2.7 giros. Un abogado piensa que las máquinas en Casino World están arregladas y observa un pago de una vez cada 12.4 giros en más de 36 máquinas. Para $\alpha = 0.01$, ¿está en lo correcto el abogado al concluir que las máquinas tienen una frecuencia de pago más baja?

capítulo 9

PRUEBA DE HIPÓTESIS: PRUEBAS DE DOS MUESTRAS

Objetivos

- Aprender a utilizar muestras de dos poblaciones para probar hipótesis acerca de cómo se relacionan esas poblaciones
- Aprender cómo la prueba de hipótesis para las diferencias entre medias de población toman diferentes formas, dependiendo de si las muestras son grandes o pequeñas
- Diferenciar muestras independientes y muestras dependientes cuando se comparan dos medias
- Aprender cómo reducir una prueba de hipótesis para la diferencia de medias, desde muestras dependientes hasta una prueba acerca de una sola media
- Aprender cómo probar hipótesis que comparan las proporciones de dos poblaciones con el mismo atributo de interés
- Entender cómo se pueden usar los valores P en las pruebas de hipótesis
- Conocer el tipo de resultados que producen los paquetes de software para estadística en pruebas de hipótesis

Contenido del capítulo

9.1 Prueba de hipótesis para diferencias entre medias y proporciones 360
9.2 Pruebas para diferencias entre medias: muestras grandes 362
9.3 Pruebas para diferencias entre medias: muestras pequeñas 366
9.4 Prueba de diferencias entre medias con muestras dependientes 372
9.5 Pruebas para diferencias entre proporciones: muestras grandes 378
9.6 Valor P: otra manera de ver las pruebas de hipótesis 386
9.7 Uso de computadoras para pruebas de hipótesis 390
- Estadística en el trabajo 392
- Ejercicio de base de datos computacional 392
- Del libro de texto al mundo real 394
- Términos introducidos en el capítulo 9 395
- Ecuaciones introducidas en el capítulo 9 395
- Ejercicios de repaso 396

Un fabricante de computadoras personales tiene un gran número de empleados de la comunidad de habla hispana. Con el fin de mejorar la productividad de su fuerza de trabajo, el fabricante desea aumentar la sensibilidad de sus administradores en cuanto a las necesidades de este grupo minoritario. Primero, programó varias sesiones de preguntas y respuestas con los líderes de la comunidad en cuestión. Después, diseñó un programa de una serie de sesiones formales en un salón de clases para promover el contacto entre sus administradores, psicólogos y sociólogos profesionales. El nuevo programa es mucho más caro que el anterior y el presidente de la compañía desea saber si el gasto ha tenido como resultado una mayor sensibilización. En este capítulo mostraremos cómo probar si estos dos métodos han tenido, en esencia, los mismos efectos sobre la sensibilidad de los administradores o si el gasto hecho en el nuevo programa queda justificado por sus resultados mejorados. ■

9.1 Prueba de hipótesis para diferencias entre medias y proporciones

Comparación de dos poblaciones

En muchas situaciones de toma de decisiones, las personas necesitan determinar si los parámetros de dos poblaciones son iguales o diferentes. Una empresa, por ejemplo, puede tener la intención de probar si sus empleadas reciben un salario menor que el de sus empleados varones por realizar el mismo trabajo. Un director de capacitación puede querer determinar si la proporción de empleados que están listos para ascensos en una dependencia gubernamental es diferente de la proporción en otra. Un fabricante de medicinas puede tener la necesidad de saber si un nuevo medicamento ocasiona una reacción en un grupo de animales para experimentación y otra reacción distinta en otro grupo.

En cada uno de estos ejemplos, los tomadores de decisiones están interesados en los parámetros de dos poblaciones; no están tan preocupados por el valor real de los parámetros como de la *relación entre* sus valores; es decir, cuáles son las diferencias. ¿Las empleadas *ganan menos* que los empleados por hacer el mismo trabajo? ¿*Es* la proporción de empleados susceptibles de promoción de una dependencia diferente a la proporción de otra? ¿Un grupo de animales para experimentación *reacciona* de manera diferente que otro grupo? En este capítulo introduciremos métodos mediante los cuales se puede dar respuesta a estas preguntas, utilizando procedimientos de pruebas de dos muestras.

Distribución de muestreo para la diferencia entre dos parámetros de población: conceptos básicos

En el capítulo 6 introdujimos el concepto de distribución de muestreo o muestral para la media como la base del trabajo que podemos hacer en estimación y pruebas de hipótesis. Para ver un resumen de la distribución de muestreo de la media, consulte la figura 6-2.

Derivación de la distribución de muestreo de la diferencia entre medias muestrales

Debido a que ahora deseamos estudiar dos poblaciones, no nada más una, la distribución de muestreo que nos interesa es la *distribución muestral de la diferencia entre medias muestrales*. La figura 9-1 puede ayudarnos a conceptualizar esta distribución de muestreo particular. En la parte superior de la figura se presentan dos poblaciones, identificadas como población 1 y población 2. Éstas tienen medias μ_1 y μ_2 y desviaciones estándar σ_1 y σ_2, respectivamente. Debajo de cada población se observa la distribución muestral de la media para la población correspondiente. En la parte inferior de la figura se encuentra la distribución muestral de la diferencia entre las medias muestrales.

Las dos distribuciones muestrales de la media teóricas de la figura 9-1 están construidas a partir de todas las muestras posibles de un tamaño dado que pueden obtenerse de la distribución de la po-

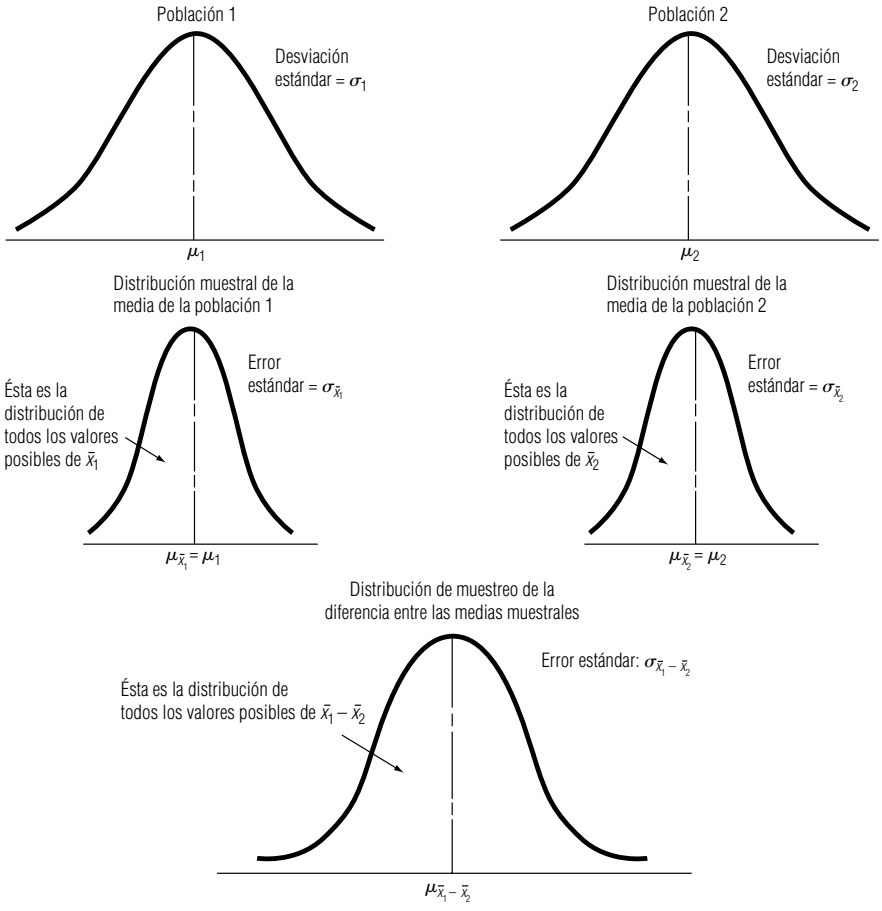


FIGURA 9-1

Conceptos básicos de distribuciones de población, distribuciones muestrales de la media y distribuciones muestrales de la diferencia entre las medias de las muestras

blación correspondiente. Ahora bien, suponga que tomamos una muestra aleatoria de la distribución de la población 1 y otra muestra aleatoria de la distribución de la población 2. Si luego restamos las dos medias de las muestras, obtenemos:

$$\bar{x}_1 - \bar{x}_2 \leftarrow \text{Diferencia entre las medias de las muestras}$$

Esta diferencia será positiva si $\bar{x}_1$ es mayor que $\bar{x}_2$, y negativa si $\bar{x}_2$ es mayor que $\bar{x}_1$. Al construir una distribución de *todas* las diferencias posibles de las muestras, $\bar{x}_1 - \bar{x}_2$, terminamos con la distribución muestral de la diferencia entre las medias de las muestras, que se ilustran la parte inferior de la figura 9-1.

Parámetros de esta distribución de muestreo

La *media de la distribución muestral de la diferencia entre las medias muestrales* se representan por $\mu_{\bar{x}_1 - \bar{x}_2}$ y es igual a $\mu_{\bar{x}_1} - \mu_{\bar{x}_2}$, que, como vimos en el capítulo 6, es igual que $\mu_1 - \mu_2$. Si $\mu_1 = \mu_2$, entonces $\mu_{\bar{x}_1} - \mu_{\bar{x}_2} = 0$.

La desviación estándar de la distribución de las diferencias entre las medias de las muestras se conoce como *error estándar de la diferencia entre dos medias* y se calcula con la siguiente fórmula:

Error estándar de la diferencia entre dos medias

Varianza de la población 1 — Varianza de la población 2

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \qquad [9\text{-}1]$$

Tamaño de la muestra de la población 1 — Tamaño de la muestra de la población 2

Cómo estimar el error estándar de esta distribución de muestreo

Si *no* conocemos las dos desviaciones estándar de la población, podemos *estimar* el error estándar de la diferencia entre dos medias. Podemos utilizar el mismo método de estimación del error estándar que hemos usado, haciendo que las desviaciones estándar de la muestra estimen las desviaciones estándar de la población de la siguiente manera:

$$\hat{\sigma} = s \quad \leftarrow \text{Desviación estándar de la muestra} \qquad [7\text{-}1]$$

Entonces, la fórmula para el error estándar estimado de la diferencia entre dos medias es

Error estándar estimado de la diferencia entre dos medias

Varianza estimada de la población 1 — Varianza estimada de la población 2

$$\hat{\sigma}_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}} \qquad [9\text{-}2]$$

Como veremos en las siguientes secciones, dependiendo de los tamaños de muestra, utilizaremos diferentes estimaciones para $\hat{\sigma}_1$ y $\hat{\sigma}_2$ en la ecuación 9-2.

9.2 Pruebas para diferencias entre medias: muestras grandes

Cuando ambos tamaños de muestra son mayores que 30, este ejemplo ilustra cómo hacer una prueba de dos colas de una hipótesis acerca de la diferencia entre dos medias. A un especialista en estadística que estudia el desarrollo de recursos humanos se le pide que determine si los salarios por hora de los obreros semicalificados son los mismos en dos ciudades distintas. El resultado de esta investigación se presenta en la tabla 9-1. Suponga que la empresa desea probar la hipótesis al nivel 0.05 de que no hay diferencia entre los salarios por hora de los trabajadores semicalificados de las dos ciudades:

Paso 1: Establezca su hipótesis, tipo de prueba y nivel de significancia

$$H_0: \mu_1 = \mu_2 \quad \leftarrow \text{Hipótesis nula: no hay diferencia}$$
$$H_1: \mu_1 \neq \mu_2 \quad \leftarrow \text{Hipótesis alternativa: existe diferencia}$$
$$\alpha = 0.05 \quad \leftarrow \text{Nivel de significancia para probar esta hipótesis}$$

Como la compañía sólo está interesada en saber si las medias son o no iguales, ésta es una prueba de dos colas.

Paso 2: Escoja la distribución apropiada y encuentre el valor crítico

Podemos ilustrar esta prueba de hipótesis gráficamente. En la figura 9-2, el nivel de significancia 0.05 corresponde a las dos áreas señaladas, cada una contiene una fracción de 0.025 del área. La región de aceptación contiene dos áreas iguales, cada una de 0.475 del área total. Como ambas muestras son grandes, podemos utilizar la distribución normal. En la tabla 1 del apéndice determinamos que el valor crítico de *z* para 0.475 del área bajo la curva es 1.96.

No se conocen las desviaciones estándar de las dos poblaciones. Por consiguiente, el primer paso es estimarlas de la siguiente manera:

$$\hat{\sigma}_1 = s_1 \qquad \hat{\sigma}_2 = s_2 \qquad [7\text{-}1]$$
$$= \$0.40 \qquad = \$0.60$$

Tabla 9-1
Datos de una encuesta con muestreo de salarios por hora

Ciudad	Salarios medios por hora de la muestra	Desviación estándar de la muestra	Tamaño de la muestra
Apex	$8.95	$0.40	200
Eden	9.10	0.60	175

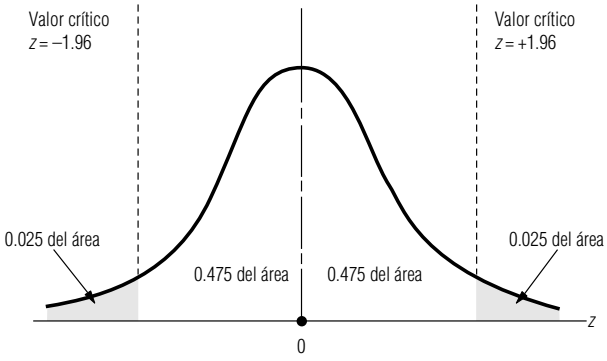


FIGURA 9-2

Prueba de hipótesis de dos colas de la diferencia entre dos medias a un nivel de significancia de 0.05

Paso 3: Calcule el error estándar y estandarice el estadístico de la muestra

Ahora podemos determinar el error estándar estimado de la diferencia entre las dos medias:

$$\hat{\sigma}_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}} \qquad [9\text{-}2]$$

$$= \sqrt{\frac{(0.40)^2}{200} + \frac{(0.60)^2}{175}}$$

$$= \sqrt{0.00286}$$

$$= \$0.053 \leftarrow \text{Error estándar estimado}$$

A continuación, estandarizamos la diferencia de las medias de las muestras, $\bar{x}_1 - \bar{x}_2$. Primero, calculamos $(\mu_1 - \mu_2)_{H_0}$, la diferencia hipotética de las medias de las poblaciones. Luego dividimos entre $\hat{\sigma}_{\bar{x}_1 - \bar{x}_2}$, el error estándar estimado de la diferencia entre las medias muestrales.

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)_{H_0}}{\hat{\sigma}_{\bar{x}_1-\bar{x}_2}}$$

$$= \frac{(8.95 - 9.10) - 0}{0.053}$$

$$= -2.83$$

Paso 4: Grafique la distribución y señale el valor de la muestra y los valores críticos

Señalamos la diferencia estandarizada en una gráfica de la distribución de muestreo y la comparamos con el valor crítico, como se ve en la figura 9-3, que señala que la diferencia estándar entre las dos medias de las muestras se encuentra fuera de la región de aceptación. Así, rechazamos la hipótesis nula de que no hay diferencia y llegamos a la conclusión de que las medias de las poblaciones (los salarios de obreros semicalificados de las dos ciudades) son diferentes.

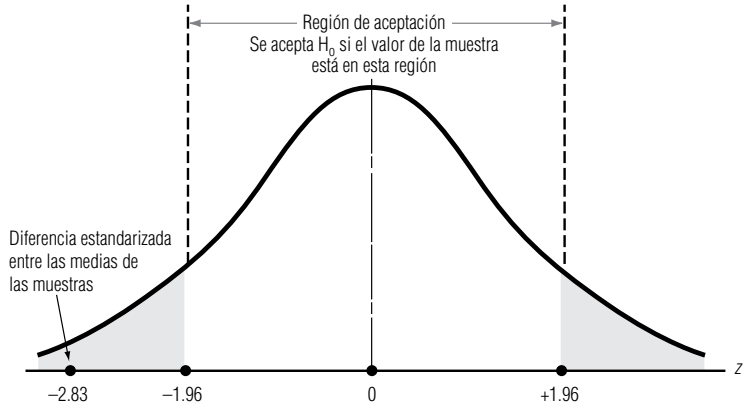


FIGURA 9-3

Prueba de hipótesis de dos colas de la diferencia entre dos medias a un nivel de significancia de 0.05, que indica la región de aceptación y la diferencia estandarizada entre las medias de las muestras

Paso 5: Interprete el resultado

Prueba de la diferencia entre medias cuando $\mu_1 - \mu_2 \neq 0$

En este ejemplo, y en la mayoría de los que veremos, se probará si dos poblaciones tienen la misma media. Si esto ocurre $(\mu_1 - \mu_2)_{H_0}$, la diferencia hipotetizada entre las dos medias es cero. Sin embargo, podemos investigar si los salarios promedio son *alrededor* de 10 centavos por hora más *bajos* en la ciudad de Apex que en Eden. En ese caso, nuestras hipótesis serían:

H_0: $\mu_1 = \mu_2 - 0.10$ ← Hipótesis nula: los salarios son \$0.10 más bajos en Apex que en Eden

H_1: $\mu_1 \neq \mu_2 - 0.10$ ← Hipótesis alternativa: los salarios no son \$0.10 más bajos en Apex que en Eden

En este caso, la diferencia hipotética entre las dos medias sería $(\mu_1 - \mu_2)_{H_0} = -0.10$, y la diferencia estandarizada entre las medias de las muestras sería:

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)_{H_0}}{\hat{\sigma}_{\bar{x}_1 - \bar{x}_2}}$$

$$= \frac{(0.95 - 9.10) - (-0.10)}{0.053}$$

$$= -0.94$$

De acuerdo con este resultado, no rechazaríamos la hipótesis nula.

Pruebas de una cola de la diferencia entre medias

Aunque el ejemplo fue de una prueba de dos colas, también podemos realizar pruebas de una cola de la diferencia entre dos medias de poblaciones. Los resultados de pruebas de una cola son conceptualmente parecidos a las pruebas de una cola de una sola media que analizamos en el capítulo 8. Por ejemplo, si hubiéramos querido probar si los salarios en Apex son *menores que* en Eden (o, de manera equivalente, si los salarios en Eden son *mayores que* en Apex), nuestras hipótesis hubieran sido:

H_0: $\mu_1 = \mu_2$ ← Hipótesis nula: los salarios son iguales en Apex y Eden

H_1: $\mu_1 < \mu_2$ ← Hipótesis alternativa: los salarios son más bajo en Apex que en Eden

Ésta sería una prueba de una cola con $(\mu_1 - \mu_2)_{H_0} = 0$.

Por último, si hubiéramos deseado probar si los salarios en Apex son *10 centavos por hora mayores que* los salarios en Eden, entonces nuestras hipótesis hubieran sido:

H_0: $\mu_1 = \mu_2 - 0.10$ ← Hipótesis nula: los salarios son \$0.10 más bajos en Apex que en Eden

H_1: $\mu_1 < \mu_2 - 0.10$ ← Hipótesis alternativa: los salarios son más de \$0.10 más bajos en Apex que en Eden

Ésta sería una prueba de una cola con $(\mu_1 - \mu_2)_{H_0} = -0.10$.

SUGERENCIAS Y SUPOSICIONES

Sugerencia: al probar las diferencias entre dos medias, debe elegir si usa una prueba de hipótesis de una cola o de dos colas. Si la prueba se refiere a si dos medias *son iguales o no son iguales*, use la prueba de dos colas que medirá si una media es diferente de la otra (mayor o menor). Si la prueba se refiere a si *una media es significativamente mayor o significativamente menor que la otra*, una prueba de una cola es apropiada.

Ejercicios 9.2

Ejercicios de autoevaluación

EA 9-1 Se recolectaron dos muestras independientes de observaciones. Para la primera muestra de 60 elementos, la media fue 86 y la desviación estándar 6. La segunda muestra de 75 elementos tenía una media de 82 y una desviación estándar de 9.

a) Calcule el error estándar estimado de la diferencia entre las dos medias.

b) Con $\alpha = 0.01$, pruebe si es razonable que se considere que las dos muestras vienen de poblaciones con la misma media.

EA 9-2 En 1993, el Consejo de Estándares para Contabilidad Financiera (CECF) consideró una propuesta para requerir que las compañías informaran el efecto potencial de la opción de compra de acciones de los empleados sobre los ingresos por acción (IPA). Una muestra aleatoria de 41 empresas de alta tecnología (AT) reveló que la nueva propuesta reduciría el IPA en un promedio del 13.8%, con una desviación estándar del 18.9%. Una nuestra aleatoria de 35 productores de bienes de consumo (BC) mostró que la propuesta reduciría el IPA en 9.1% en promedio, con desviación estándar del 8.7%. Con base en estas muestras, ¿es razonable concluir (para $\alpha = 0.10$) que la propuesta de la CECF causaría una mayor reducción en el IPA para las empresas de alta tecnología que para los productores de bienes de consumo?

Conceptos básicos

■ **9-1** Se tomaron dos muestras independientes. Para la primera de 42 elementos, la media fue 32.3 y la varianza 9. La segunda muestra de 57 elementos tenía media de 34 y varianza de 16.

a) Calcule el error estándar estimado de la diferencia entre las dos medias.

b) Con $\alpha = 0.05$, pruebe si existe suficiente evidencia para mostrar que la segunda población tiene una media mayor.

Aplicaciones

■ **9-2** Block, una compañía fabricante de chips para computadoras, está en proceso de decidir si sustituye su línea de ensamble semiautomática por otra completamente automatizada. Block ha reunido algunos datos de pruebas preliminares acerca de la producción de chips por hora que se resumen en la tabla siguiente y desea saber si debe actualizar su línea de ensamble. Establezca (y pruebe con $\alpha = 0.02$) las hipótesis apropiadas para ayudar a Block a tomar una decisión.

	$\bar{x}$	s	n
Línea semiautomática	198	32	150
Línea automática	206	29	200

■ **9-3** Dos laboratorios de investigación han producido, de manera independiente, medicamentos que alivian las molestias de la artritis. El primer medicamento fue probado en un grupo de 90 personas que sufren la enfermedad y produjo un promedio de 8.5 horas de alivio, con desviación estándar de 1.8 horas. El segundo fue probado en 80 artríticos y produjo una media de 7.9 horas de alivio, con desviación estándar de 2.1 horas. A un nivel de significancia de 0.05, ¿el segundo medicamento proporciona un periodo de alivio significativamente más corto?

■ **9-4** El 1 de enero de 1996 se tomó una muestra de 32 fondos mutualistas de la bolsa de valores, y se encontró que la tasa promedio de rendimiento anual durante los 30 días anteriores fue del 3.23%, con una desviación estándar de la muestra del 0.51%. Un año antes, una muestra de 38 fondos mutualistas indicó una tasa promedio de rendimiento del 4.36%, con una desviación estándar de la muestra del 0.84%. ¿Es razonable llegar a la conclusión (a un nivel $\alpha = 0.05$) de que las tasas de interés del mercado de dinero declinaron durante 1995?

■ **9-5** En septiembre de 1995, la Confederación Automovilística de las Carolinas investigó al azar a 75 gasolineras en Carolina del Norte y Carolina del Sur y determinó que el precio promedio de la gasolina regular sin plomo en las bombas de autoservicio fue \$1.059, con una desviación estándar de 3.9 centavos. Tres meses después, en otra investigación aleatoria de 50 gasolineras, se encontró un precio promedio de \$1.089, con una desviación estándar de 6.8 centavos. A un nivel $\alpha = 0.02$, ¿cambió significativamente el precio de la gasolina regular sin plomo en estos dos estados durante estos tres meses?

■ **9-6** A pesar de la Ley de Pagos Igualitarios de 1963, en 1993 todavía parecía que los hombres ganaban más que las mujeres en trabajos similares. En una muestra aleatoria de 38 operadores varones de máquinas-herramienta se encontró que el salario medio por hora era \$11.38, con una desviación estándar de la muestra de \$1.84. Se tomó una muestra aleatoria de 45 operadoras de máquinas-herramienta y se obtuvo un salario medio por hora de \$8.42, con desviación estándar de la muestra de \$1.31. Según estas dos muestras, ¿es razonable llegar a la conclusión (a un nivel $\alpha = 0.01$) de que los operadores ganan más de \$2.00 por hora que las operadoras?

■ **9-7** La tienda de descuento BullsEye está orgullosa del servicio que presta a sus clientes. La tienda espera que toda la cadena esté dando el mismo nivel de servicio de costa a costa, así que encuestaron algunos clientes. En el sureste, una muestra aleatoria de 97 clientes dio una calificación de la satisfacción global promedio de 8.8 sobre 10 puntos con desviación estándar de la muestra de 0.7. En el noreste, la muestra aleatoria de 84 clientes dio una calificación promedio de 9.0 y la desviación estándar de la muestra fue 0.6. ¿Puede concluir BullsEye, con $\alpha = 0.05$, que los niveles de satisfacción de los clientes en los dos mercado son significativamente diferentes?

Soluciones a los ejercicios de autoevaluación

EA 9-1 $s_1 = 6 \quad n_1 = 60 \quad \bar{x}_1 = 86 \quad s_2 = 9 \quad n_2 = 75 \quad \bar{x}_2 = 82$

a) $\hat{\sigma}_{\bar{x}_1 - \bar{x}_2} = \sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}} = \sqrt{\dfrac{36}{60} + \dfrac{81}{75}} = 1.296$

b) $H_0: \mu_1 = \mu_2 \quad H_1: \mu_1 \neq \mu_2 \quad \alpha = 0.01$

Los límites de la región de aceptación son $z = \pm 2.58$, o

$$\bar{x}_1 - \bar{x}_2 = 0 \pm z\hat{\sigma}_{\bar{x}_1 - \bar{x}_2} = \pm 2.58(1.296) = \pm 3.344$$

Debido a que el valor z observado $= \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)_{H_0}}{\hat{\sigma}_{\bar{x}_1 - \bar{x}_2}}$

$$= \frac{(86 - 82) - 0}{1.296}$$

$= 3.09 > 2.58$ (o $\bar{x}_1 - \bar{x}_2 = 86 - 82 = 4 > 3.344$), se rechaza H_0. Es razonable concluir que las dos muestras vienen de poblaciones diferentes.

EA 9-2 Muestra 1 (empresas de AT): $s_1 = 18.9 \quad n_1 = 41 \quad \bar{x}_1 = 13.8$

Muestra 2 (productores de BC): $s_2 = 8.7 \quad n_2 = 35 \quad \bar{x}_2 = 9.1$

$H_0: \mu_1 = \mu_2 \quad H_1: \mu_1 > \mu_2 \quad \alpha = 0.10$

$$\hat{\sigma}_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{(18.9)^2}{41} + \frac{(8.7)^2}{35}} = 3.298\%$$

El límite superior de la región de aceptación es $z = 1.28$, o

$$\bar{x}_1 - \bar{x}_2 = 0 + z\hat{\sigma}_{\bar{x}_1 - \bar{x}_2} = 1.28(3.298) = 4.221\%$$

Como el valor z observado $= \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)H_0}{\hat{\sigma}_{\bar{x}_1 - \bar{x}_2}} = \dfrac{(13.8 - 9.1) - 0}{3.298} = 1.43 > 1.28$ (o $\bar{x}_1 - \bar{x}_2 = 4.7 > 4.221$), se rechaza H_0 y se concluye que la propuesta de la CECF causará una reducción significativamente mayor en el IPA de las empresas de alta tecnología.

9.3 Pruebas para diferencias entre medias: muestras pequeñas

Cuando los tamaños de muestra son pequeños, debemos hacer dos cambios técnicos en el procedimiento para probar las diferencias entre medias. El primero tiene que ver con la forma en que calculamos el error estándar estimado de la diferencia entre dos medias muestrales. El segundo le recordará lo que hicimos en el capítulo 8 con las pruebas de muestras pequeñas de una sola media. De nuevo, basaremos nuestras pruebas de muestra pequeña en la distribución *t*, más que en la distribución normal. Para explorar los detalles de estos cambios, regresemos al ejemplo introductorio del presente capítulo, concerniente a la sensibilización de los administradores de una fábrica de computadoras personales con respecto a las necesidades de sus trabajadores de habla hispana.

Recuerde que la compañía ha estado investigando dos programas educativos para aumentar la sensibilidad de sus directores. El programa original consistía en varias sesiones informales de pre-

guntas y respuestas con los líderes de la comunidad hispana. En años recientes, se ha desarrollado un programa que implica clases formales con psicólogos y sociólogos profesionales. El nuevo programa es considerablemente más caro, y el presidente de la empresa desea saber, a un nivel de 0.05 de significancia, si este gasto ha aumentado la sensibilidad de sus administradores. Probemos lo siguiente:

Paso 1: Establezca su hipótesis, tipo de prueba y nivel de significancia

H_0: $\mu_1 = \mu_2$ ← Hipótesis nula: no hay diferencia en los niveles de sensibilidad logrados por los dos programas

H_1: $\mu_1 > \mu_2$ ← Hipótesis alternativa: el nuevo programa da resultados en niveles de sensibilidad más altos

$\alpha = 0.05$ ← Nivel de significancia para probar esta hipótesis

La tabla 9-2 contiene los datos obtenidos de una muestra de administradores capacitados en ambos programas. Debido a que sólo se dispone de un número limitado de datos para los dos programas, las desviaciones estándar de la población se estiman a partir de los datos. El nivel de sensibilidad se mide como porcentaje en una escala psicométrica estándar.

La compañía desea probar si la sensibilidad adquirida después de tomar el nuevo programa es *significativamente mayor* que la adquirida con el programa informal anterior. Para rechazar la hipótesis nula (un resultado que desea la compañía), la diferencia observada de las medias de las muestras necesitaría encontrarse suficientemente alejada en la cola *derecha* de la distribución. Entonces aceptaríamos la hipótesis alternativa de que el nuevo programa lleva a mayores niveles de sensibilidad y que los gastos adicionales hechos en este programa están justificados.

Posponga el paso 2 hasta saber cuántos grados de libertad usar

El segundo paso de nuestro proceso de cinco pasos para la prueba de hipótesis requiere elegir la distribución apropiada y encontrar el valor crítico. Recuerde que en el primer párrafo de esta sección afirmamos que la prueba estaría basada en una distribución *t*, pero aún no sabemos cuál debemos utilizar. *¿Cuántos grados de libertad existen?* La respuesta se hará más evidente después de ver cómo se calcula el error estándar estimado.

Nuestra primera tarea al efectuar la prueba consiste en calcular el error estándar de la diferencia entre las dos medias. Como no se conocen las desviaciones estándar de las poblaciones, debemos utilizar la ecuación 9-2.

$$\hat{\sigma}_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}} \qquad [9\text{-}2]$$

En el ejemplo anterior, donde los tamaños de muestra eran grandes (ambos mayores que 30), usamos la ecuación 7-1 y estimamos $\hat{\sigma}_1^2$ mediante s_1^2, y $\hat{\sigma}_2^2$ mediante s_2^2. En este caso, con muestras pequeñas, el procedimiento no es adecuado. Si podemos suponer que las varianzas de población desconocidas son iguales (esta suposición puede probarse con un método analizado en la sección 6 del capítulo 11), entonces podemos seguir adelante. Si no podemos suponer que $\sigma_1^2 = \sigma_2^2$, entonces el problema está más allá del nivel de este libro.

Estimación de σ^2 con muestras pequeñas

Supongamos, por el momento, que $\sigma_1^2 = \sigma_2^2$, ¿de qué manera podemos estimar la varianza común σ^2? Si utilizamos s_1^2 o s_2^2, obtenemos un estimador imparcial de σ^2, pero no usamos toda la información que tenemos disponible, ya que se ignora una de las muestras. En su lugar, usamos un promedio ponderado de s_1^2 y s_2^2, en el cual los pesos son el número de grados de libertad de cada muestra. A este promedio ponderado se le conoce como "estimación conjunta" de σ^2 y está dado por:

Estimación conjunta de σ^2

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \qquad [9\text{-}3]$$

Para esta prueba, tenemos $n_1 + n_2 - 2$ grados de libertad

Como tenemos que usar las varianzas de la muestra para estimar el valor desconocido de σ^2, la prueba estará basada en la distribución *t*. Este caso es igual a probar una sola media de tamaño *n*, cuando no conocemos el valor de σ^2. Ahí utilizamos una distribución *t* con $n - 1$ grados de libertad, debido

Tabla 9-3

Datos de las muestras de dos programas de sensibilización

Programa muestreado	Sensibilidad media después del programa	Número de administradores observados	Desviación estándar estimada de la sensibilidad después del programa
Formal	92%	12	15%
Informal	84	15	19

a que una vez que conocemos la media de la muestra sólo $n - 1$ observaciones se pueden especificar libremente. (Tal vez desee repasar el análisis de grados de libertad en el capítulo 7.) Ahora, tenemos $n_1 - 1$ grados de libertad en la primera muestra y $n_2 - 1$ grados de libertad en la segunda, de modo que cuando las unimos para estimar α^2, obtenemos $n_1 + n_2 - 2$ grados de libertad. Así, la distribución de muestreo apropiada para la prueba de los dos programas de sensibilidad es la distribución t con $12 + 15 - 2 = 25$ grados de libertad. Debido a que estamos efectuando una prueba de cola superior a un nivel de significancia de 0.05, el valor crítico de t es 1.708, de acuerdo con la tabla 2 del apéndice.

Regreso al paso 2: seleccione la distribución apropiada y encuentre el valor crítico

La figura 9-4 ilustra la prueba de hipótesis, ahora que ya tenemos el valor crítico para la misma. La región sombreada que se encuentra a la derecha de la distribución representa el nivel 0.05 de significancia de la prueba.

En el paso 3, insertamos la fórmula para s_p^2 de la ecuación 9-3 en la ecuación 9-2 y simplificamos el resultado para obtener una ecuación para el error estándar estimado de $\bar{x}_1 - \bar{x}_2$:

Inicio del paso 3: calcule el error estándar

Error estándar estimado de la diferencia entre dos medias muestrales, con muestras pequeñas y varianzas de población iguales

$$\hat{\sigma}_{\bar{x}_1 - \bar{x}_2} = s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \qquad [9\text{-}4]$$

Aplicando estos resultados a nuestro ejemplo de sensibilidad:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \qquad [9\text{-}3]$$

$$= \frac{(12 - 1)(15)^2 + (15 - 1)(19)^2}{12 + 15 - 2}$$

$$= \frac{11(225) + 14(361)}{25}$$

$$= 301.160$$

Sacando la raíz cuadrada en ambos lados, obtenemos $s_p = \sqrt{301.160}$, o 17.354, y por tanto:

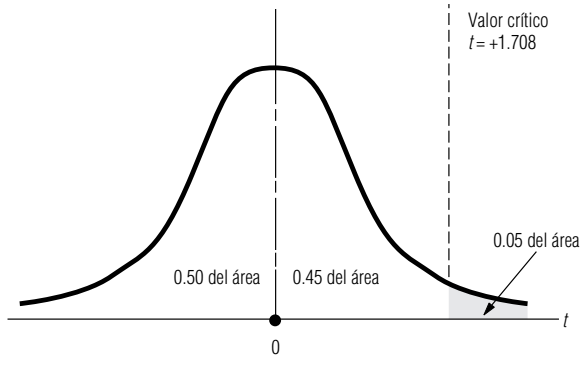


FIGURA 9-4

Prueba de hipótesis de cola derecha para la diferencia entre dos medias a un nivel de significancia de 0.05

$$\hat{\sigma}_{\bar{x}_1-\bar{x}_2} = s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \qquad [9\text{-}4]$$

$$= 17.354\sqrt{\frac{1}{12}+\frac{1}{15}}$$

$$= 17.354(0.387)$$

$$= 6.721 \leftarrow \text{Error estándar estimado de la diferencia}$$

Conclusión del paso 3: estandarice el estadístico de la muestra

A continuación estandarizamos la diferencia de las medias de las muestras, $\bar{x}_1 - \bar{x}_2$. Primero, restamos $(\mu_1 - \mu_2)_{H_0}$, la diferencia hipotética de las medias muestrales. Luego dividimos entre $\hat{\sigma}_{\bar{x}_1-\bar{x}_2}$, el error estándar estimado de la diferencia entre las medias muestrales.

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)_{H_0}}{\hat{\sigma}_{\bar{x}_1-\bar{x}_2}}$$

$$= \frac{(92 - 84) - 0}{6.721}$$

$$= 1.19$$

Debido a que nuestra prueba de hipótesis está basada en la distribución *t*, utilizamos *t* para representar al estadístico estandarizado.

Paso 4: Grafique un diagrama de la distribución y señale el valor de la muestra y el valor crítico

Paso 5: Interprete el resultado

Después, señalamos la diferencia estándar en una gráfica de la distribución de muestreo y la comparamos con el valor crítico $t = 1.708$, como se ilustra en la figura 9-5. En ella podemos apreciar que la diferencia estandarizada entre las dos medias de las muestras se encuentra dentro de la región de aceptación. Así, aceptamos la hipótesis nula de que no existe una diferencia significativa entre los niveles de sensibilidad logrados por los dos programas. Los gastos de la empresa en el programa formal no han producido un incremento significativo en el nivel de sensibilidad de sus administradores.

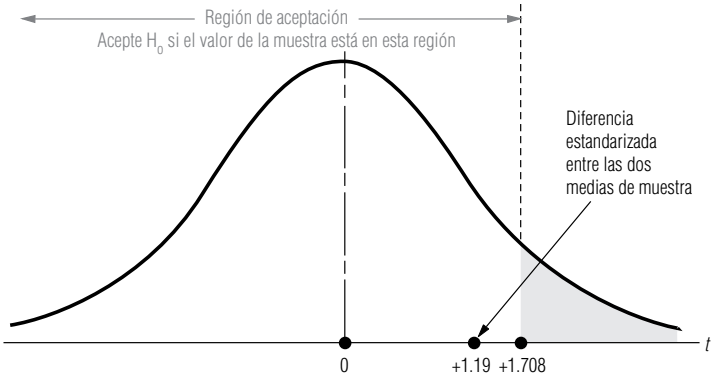


FIGURA 9-5

Prueba de una cola de la diferencia entre dos medias al nivel 0.05 de significancia; indica la región de aceptación y la diferencia estandarizada entre las medias muestrales

SUGERENCIAS Y SUPOSICIONES

Sugerencia: debido a que los tamaños de muestra son pequeños (menos de 30) y no se conocen las desviaciones estándar de las poblaciones, será adecuado usar la distribución *t*. Al igual que en la prueba *t* de una sola muestra que se estudió, en este caso también es necesario determinar los grados de libertad. En la prueba de una muestra, los grados de libertad eran el tamaño de la muestra menos uno. Ahora, como se usan dos muestras, los grados de libertad correctos son el tamaño de la primera muestra menos uno más el tamaño de la segunda muestra menos uno: $n_1 + n_2 - 2$. Suposición: se parte del supuesto de que las varianzas de las dos poblaciones son iguales. Si no es así, no se puede realizar esta prueba usando los métodos descritos. Advertencia: para usar el método explicado en esta sección, las dos muestras (una de cada población) deben haberse elegido de manera que sean independientes entre sí.

Ejercicios 9.3

Ejercicios de autoevaluación

EA 9-3 Una organización de investigación de mercados selecciona varios modelos de automóviles cada año y evalúa su eficiencia en el consumo de combustible. Este año, en el análisis de dos modelos subcompactos similares de dos fabricantes distintos, el millaje promedio de 12 autos de la marca A fue 27.2 millas por galón, y la desviación estándar fue 3.8 mpg. Los 9 autos de la marca B que se probaron promediaron 32.1 mpg con desviación estándar de 4.3 mpg. Para $\alpha = 0.01$, ¿se puede concluir que la marca A da un millaje promedio menor que la marca B?

EA 9-4 Connie Rodriguez, la decana de estudiantes en el Midstate College, se pregunta cuál será la distribución de calificaciones en la escuela. Ha oído quejas de que el promedio general en la escuela de administración está cerca de 0.25 más abajo que en las universidades de artes y ciencias. Un muestreo aleatorio rápido produjo los siguientes promedios generales.

Administración:	2.86	2.77	3.18	2.80	3.14	2.87	3.19	3.24	2.91	3.00	2.83		
Artes y ciencias:	3.35	3.32	3.36	3.63	3.41	3.37	3.45	3.43	3.44	3.17	3.26	3.18	3.41

¿Indican estos datos que existe una base para las quejas? Establezca y pruebe las hipótesis adecuadas para $\alpha = 0.02$.

Aplicaciones

■ **9-8** Una organización de crédito y seguros ha desarrollado un nuevo método de alta tecnología para capacitar al nuevo personal de ventas. La compañía obtuvo una muestra de 16 empleados capacitados de la manera original y encontró ventas diarias promedio de $688 con desviación estándar de la muestra de $32.63. También tomaron una muestra de 11 empleados capacitados con el método nuevo y encontraron un promedio de ventas diarias de $706 con desviación estándar de la muestra de $24.84. Para $\alpha = 0.05$, ¿puede la compañía concluir que el promedio diario de ventas aumenta con el nuevo plan?

■ **9-9** Una empresa grande de corretaje de acciones desea determinar qué tanto éxito han tenido sus nuevos ejecutivos de cuenta en la consecución de clientes. Después de terminar su capacitación, los nuevos ejecutivos pasan varias semanas haciendo llamadas a posibles clientes, tratando de que los prospectos abran cuentas con la empresa. Los datos siguientes dan el número de cuentas nuevas abiertas durante las primeras dos semanas por 10 ejecutivas y 8 ejecutivos de cuenta escogidos aleatoriamente. A un nivel de $\alpha = 0.05$, ¿parece que las mujeres son más efectivas que los hombres para conseguir nuevas cuentas?

	Número de cuentas nuevas									
Ejecutivas de cuenta	12	11	14	13	13	14	13	12	14	12
Ejecutivos de cuenta	13	10	11	12	13	12	10	12		

■ **9-10** Para celebrar su primer aniversario, Randy Nelson decidió comprar un par de aretes de diamantes para su esposa Debbie. Le enseñaron 9 pares de aretes con gemas que pesaban aproximadamente 2 quilates por par. Debido a las diferencias en color y calidad de las piedras, los precios variaban de una joya a otra. El precio promedio fue $2,990, con una desviación estándar de la muestra de $370. Además le enseñaron 6 pares con piedras en forma de gota, también con un peso aproximado de 2 quilates por par. Estos pendientes tenían un precio promedio de $3,065 con desviación estándar de $805. Con base en esta evidencia, ¿puede Randy llegar a la conclusión (a un nivel de significancia de 0.05) de que los diamantes con forma de gota cuestan más, en promedio, que los otros?

■ **9-11** Una muestra de tasas hipotecarias convencionales a 30 años tomadas al azar en 11 bancos de California produjo una tasa media del 7.61% y una desviación estándar del 0.39%. Una muestra parecida tomada aleatoriamente en ocho bancos de Pennsylvania tuvo una tasa media del 7.43%, con desviación estándar del 0.56%. ¿Estas muestras proporcionan evidencia para llegar a la conclusión (a un nivel $\alpha = 0.10$) de que las tasas de hipotecas convencionales de California y Pennsylvania provienen de poblaciones con medias distintas?

■ **9-12** Debido a que los reembolsos de impuestos se pagan con más rapidez cuando se solicitan electrónicamente, el comisionado del Servicio Interno de Contribuciones se preguntaba si los reembolsos por devolución de impuestos solicitados por correo eran menores que los solicitados electrónicamente. Observando sola-

mente los reembolsos reclamados, una muestra de 17 solicitados por correo tuvo un reembolso medio de $563 y una desviación estándar de $378. Los reembolsos promedio reclamados en 13 solicitudes electrónicas fueron de $958, con desviación estándar de la muestra de $619. A un nivel $\alpha = 0.01$, ¿estos datos apoyan la especulación del Comisionado?

■ **9-13** En la actualidad, Llantas Greatyear produce sus neumáticos en la planta de Wilmington, Carolina del Norte, con dos turnos de 12 horas. Los empleados del turno de noche planean pedir un aumento porque piensan que están produciendo más llantas por turno que el turno de día. "Como la compañía gana más durante el turno de noche, esos empleados también deben ganar más", declara el representante de ese turno. I. M. Checking, el supervisor de producción de Greatyear, selecciona al azar algunas corridas de producción diarias de cada turno con los resultados que se presentan en la tabla (en miles de llantas producidas).

Turno	Producción (en miles)								
Día	107.5	118.6	124.6	101.6	113.6	119.6	120.6	109.6	105.9
Noche	115.6	109.4	121.6	128.7	136.6	125.4	121.3	108.6	117.5

¿Indican estos datos, para $\alpha = 0.01$, que el turno de noche produce más llantas por turno?

Soluciones a los ejercicios de autoevaluación

EA 9-3 $s_A = 3.8$ $n_A = 12$ $\bar{x}_A = 27.2$ $s_B = 4.3$ $n_B = 9$ $\bar{x}_B = 32.1$

H_0: $\mu_A = \mu_B$ H_1: $\mu_A < \mu_B$ $\alpha = 0.01$

$$s_p = \sqrt{\frac{(n_A - 1)s_A^2 + (n_B - 1)s_B^2}{n_A + n_B - 2}} = \sqrt{\frac{11(3.8)^2 + 8(4.3)^2}{19}} = 4.0181 \text{ mpg}$$

El límite inferior de la región de aceptación es $t = -2.539$, o

$$\bar{x}_A - \bar{x}_B = 0 - ts_p\sqrt{\frac{1}{n_A} + \frac{1}{n_B}} = -2.539(4.0181)\sqrt{\frac{1}{12} + \frac{1}{9}}$$

$$= -4.499 \text{ mpg}$$

Como el valor observado $t = \dfrac{(\bar{x}_A - \bar{x}_B) - (\mu_A - \mu_B)_{H_0}}{s_p\sqrt{\dfrac{1}{n_A} + \dfrac{1}{n_B}}} = \dfrac{(27.2 - 32.1) - 0}{4.0181\sqrt{\dfrac{1}{12} + \dfrac{1}{9}}}$

$= -2.766 < -2.539$ (o $\bar{x}_A - \bar{x}_B = -4.9 < -4.499$), se rechaza H_0. La marca B entrega un millaje significativamente mayor que el de la marca A.

EA 9-4 Muestra 1 (administración): $s_B = 0.176$ $n_B = 11$ $\bar{x}_B = 2.98$

Muestra 2 (artes y ciencias): $s_A = 0.121$ $n_A = 13$ $\bar{x}_A = 3.368$

H_0: $\mu_B - \mu_A = -0.25$ H_1: $\mu_B - \mu_A \neq -0.25$ $\alpha = 0.02$

$$s_p = \sqrt{\frac{(n_B - 1)s_B^2 + (n_A - 1)s_A^2}{n_B + n_A - 2}} = \sqrt{\frac{10(0.176)2 + 12(0.121)^2}{22}} = 0.1485$$

Los límites de la región de aceptación son $t = \pm 2.508$, o

$$\bar{x}_B - \bar{x}_A = (\mu_B - \mu_A)_{H_0} \pm ts_p\sqrt{\frac{1}{n_B} + \frac{1}{n_A}} = -0.25$$

$$\pm 2.508(0.1485)\sqrt{\frac{1}{11} + \frac{1}{13}} = (-0.4026, -0.0974)$$

Como el valor observado $t = \dfrac{(\bar{x}_B - \bar{x}_A) - (\mu_B - \mu_A)_{H_0}}{s_p\sqrt{\dfrac{1}{n_B} + \dfrac{1}{n_A}}}$

$$= \frac{(2.980 - 3.368) - 0.25}{0.1485\sqrt{\dfrac{1}{11} + \dfrac{1}{13}}}$$

$= -2.268 > -2.508$ (o $\bar{x}_B - \bar{x}_A$ $-0.388 > -0.403$), no se rechaza H_0. El promedio general en la escuela de administración está alrededor de 0.25 abajo del de las universidades de artes y ciencias.

9.4 Prueba de diferencias entre medias con muestras dependientes

Condiciones para las cuales las muestras apareadas ayudan al análisis

En los ejemplos de las secciones 9.2 y 9.3, las muestras fueron escogidas *de manera independiente* una de otra. En el ejemplo de los salarios, se tomaron en dos ciudades distintas; en el de la sensibilidad de los administradores, se obtuvieron entre funcionarios que habían cursado dos programas de capacitación diferentes. En ciertas ocasiones, sin embargo, tiene sentido tomar muestras que no son independientes entre sí. A menudo, el uso de muestras *dependientes* (o *apareadas*) permite llevar a cabo un análisis más preciso, porque permite controlar factores externos. Con muestras dependientes, todavía se sigue el procedimiento básico adoptado en todas las pruebas de hipótesis. Las únicas diferencias consisten en que se emplea una fórmula distinta para el error estándar estimado de las diferencias muestrales y que es necesario que ambas muestras sean del mismo tamaño.

Un balneario de aguas curativas anuncia un programa de reducción de peso y afirma que el participante promedio pierde más de 17 libras. Un ejecutivo un tanto sobrado de peso está interesado en el programa, pero duda de lo que afirma el anuncio y solicita evidencia más fuerte. El balneario le permite elegir al azar los registros de 10 participantes y anotar su peso (en libras) antes y después del programa. Estos datos se presentan en la tabla 9-3. En este ejemplo tenemos dos muestras (una muestra de *antes* y una de *después*) que son claramente dependientes entre sí, pues la muestra de 10 personas se observó dos veces.

Paso 1: Establezca sus hipótesis, tipo de prueba y nivel de significancia

El ejecutivo desea probar, a un nivel de significancia del 5%, la pérdida promedio de peso anunciada de más de 17 libras. Formalmente, este problema se plantea:

H_0: $\mu_1 - \mu_2 = 17$ ← Hipótesis nula: la pérdida promedio de peso es sólo 17 libras

H_1: $\mu_1 - \mu_2 > 17$ ← Hipótesis alternativa: la pérdida promedio de peso excede 17 libras

$\alpha = 0.05$ ← Nivel de significancia

Comprensión del concepto de diferencias

Lo que en realidad nos interesa no son los pesos antes y después del tratamiento, sino en sus *diferencias*. **Conceptualmente, lo que tenemos *no son dos muestras* de peso antes y después, sino más bien *una muestra* de pérdidas de peso.** Si la población de pérdidas de peso tiene una media μ_l, podemos replantear nuestra hipótesis como:

$$H_0: \mu_l = 17$$
$$H_1: \mu_l > 17$$

Tabla 9-3

Pesos antes y después de un programa de reducción de peso (libras)

Antes	189	202	220	207	194	177	193	202	208	233
Después	170	179	203	192	172	161	174	187	186	204

Paso 2: Escoja la distribución apropiada y encuentre el valor crítico

La figura 9-6 ilustra este problema. Como deseamos saber si la pérdida media de peso *excede* 17 libras, es apropiada una prueba de cola superior. El nivel de significancia de 0.05 aparece en la figura 9-6 como el área sombreada bajo la distribución *t*. Utilizamos la distribución *t* debido a que el tamaño de la muestra es sólo 10; el número correcto de grados de libertad es 9 (10 − 1). En la tabla 2 del apéndice se da el valor critico de *t*, 1.833.

Cálculo de las diferencias por pares

Empezamos con el cálculo de las pérdidas individuales, su media y su desviación estándar, y procedemos de la misma manera que al probar hipótesis sobre una sola media. Los cálculos se muestran en la tabla 9-4.

A continuación, utilizamos la ecuación 7-1 para estimar la desviación estándar desconocida de la población:

$$\hat{\sigma} = s = 4.40 \qquad [7\text{-}1]$$

Paso 3: Calcule el error estándar y estandarice el estadístico de la muestra

y ahora podemos estimar el error estándar de la media:

$$\hat{\sigma}_{\bar{x}} = \frac{\hat{\sigma}}{\sqrt{n}} \qquad [7\text{-}6]$$

Tabla 9-4

Búsqueda de la media de pérdidas de peso y su desviación estándar

Antes	Después	Pérdida x	Pérdida al cuadrado x^2
189	170	19	361
202	179	23	529
220	203	17	289
207	192	15	225
194	172	22	484
177	161	16	256
193	174	19	361
202	187	15	225
208	186	22	484
233	204	29	841
		$\Sigma x = \mathbf{197}$	$\Sigma x^2 = \mathbf{4{,}055}$

$$\bar{x} = \frac{\Sigma x}{n} \quad [3\text{-}2] \qquad = \frac{197}{10} \qquad = 19.7$$

$$s = \sqrt{\frac{\Sigma x^2}{n-1} - \frac{n\bar{x}^2}{n-1}} \quad [3\text{-}18] \qquad = \sqrt{\frac{4{,}055}{9} - \frac{10(19.7)^2}{9}} \qquad = \sqrt{19.34} \qquad = 4.40$$

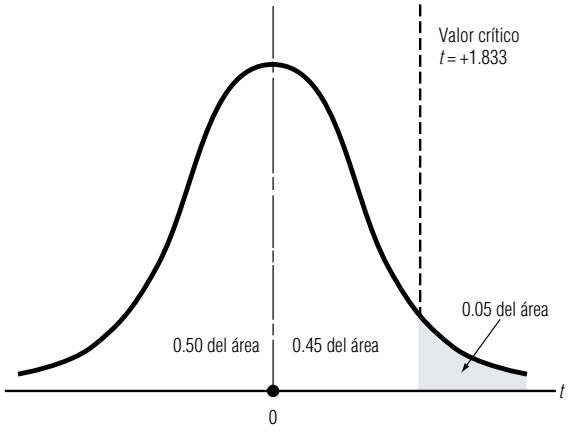


FIGURA 9-6

Prueba de hipótesis de una cola al nivel 0.05 de significancia

Paso 4: Grafique la distribución y señale el valor de la muestra y el valor crítico

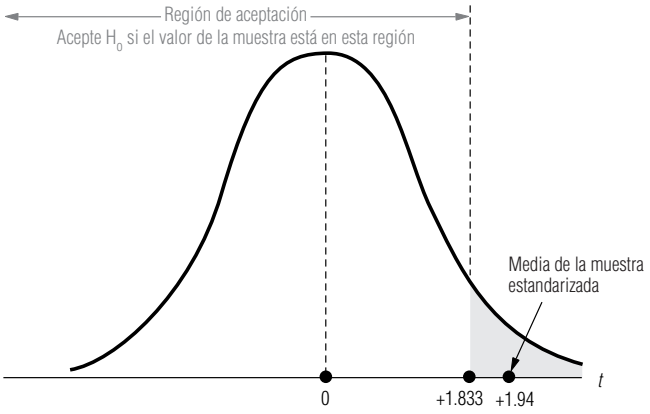


FIGURA 9-7

Prueba de hipótesis de una cola al nivel 0.05 de significancia, en la que se indica la región de aceptación y la media de la muestra estandarizada

$$= \frac{4.40}{\sqrt{10}}$$

$$= \frac{4.40}{3.16}$$

$$= 1.39 \leftarrow \text{Error estándar estimado de la media}$$

En seguida, estandarizamos la pérdida media de peso observada, $\bar{x} = 19.7$ libras, restándole μ_{H_0}, la pérdida media hipotética, y dividiendo el resultado entre $\hat{\sigma}_{\bar{x}}$, el error estándar estimado de la media:

$$t = \frac{\bar{x} - \mu_{H_0}}{\hat{\sigma}_{\bar{x}}}$$

$$= \frac{19.7 - 17}{1.9}$$

$$= 1.94$$

Debido a que nuestra prueba de hipótesis se basa en la distribución *t*, utilizamos *t* para representar el estadístico estandarizado.

Paso 5: Interprete el resultado

La figura 9-7 ilustra la localización de la pérdida media de peso en la escala estandarizada. Vemos que la media de la muestra se encuentra fuera de la región de aceptación, de modo que el ejecutivo puede rechazar la hipótesis nula y llegar a la conclusión de que la pérdida de peso anunciada con el programa es legítima.

¿En qué es distinta la prueba de diferencias por pares?

Veamos en qué esta prueba de *diferencias por pares* es distinta de una prueba de la diferencia de las medias de *dos muestras independientes*. Suponga que los datos de la tabla 9-4 representan dos muestras independientes, una de 10 individuos que *entran* al programa y *otra* distinta de 10 individuos elegidos al azar que *terminan* el programa. Las medias y varianzas de las dos muestras se dan en la tabla 9-5.

Una estimación conjunta de σ^2

Como se trata de muestras pequeñas, utilizamos la ecuación 9-3 para obtener una estimación conjunta de σ^2 y la ecuación 9-4 para estimar $\sigma_{\bar{x}_1 - \bar{x}_2}$:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \qquad [9\text{-}3]$$

$$= \frac{(10 - 1)(253.61) + (10 - 1)(201.96)}{10 + 10 - 2}$$

$$= \frac{2282.49 + 1817.64}{18}$$

$$= 227.79 \leftarrow \text{Estimación de la varianza de la población común}$$

$$\hat{\sigma}_{\bar{x}_1-\bar{x}_2} = s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \qquad [9\text{-}4]$$

$$= \sqrt{227.79}\sqrt{\frac{1}{10}+\frac{1}{10}}$$

$$= 15.09(0.45)$$

$$= 6.79 \leftarrow \text{Estimación de } \sigma_{\bar{x}_1-\bar{x}_2}$$

La prueba apropiada está basada, ahora, en la distribución *t* con 18 grados de libertad (10 + 10 − 2). Con un nivel de significancia de 0.05, el valor crítico de *t*, de la tabla 2 del apéndice, es 1.734. La diferencia observada de las medias muestrales es:

$$\bar{x}_1 - \bar{x}_2 = 202.5 - 182.8$$
$$= 19.7 \text{ libras}$$

Ahora bien, cuando estandarizamos la diferencia de las medias de las muestras para esta prueba de muestras independientes, obtenemos:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)_{H_0}}{\hat{\sigma}_{\bar{x}_1-\bar{x}_2}}$$

$$= \frac{(202.5 - 182.8) - 17}{6.79}$$

$$= 0.40$$

Con muestras independientes, H_0 no puede ser rechazada

De nuevo, debido a que nuestra prueba de hipótesis se basa en la distribución *t*, utilizamos *t* para representar el estadístico estandarizado. Al comparar la diferencia estandarizada de las medias de las muestras (0.40) con el valor crítico de *t* (1.734), vemos que el estadístico estandarizado de la muestra ya no se encuentra fuera de la región de aceptación, de modo que esta prueba *no* rechazará a H_0.

Explicación de los resultados diferentes

¿Por qué estas dos pruebas producen resultados distintos? En la prueba de muestras apareadas, la desviación estándar de la muestra de las diferencias individuales era relativamente pequeña, de manera que 19.7 libras era una cantidad significativamente mayor que la pérdida de peso hipotética de 17 libras. Sin embargo, con muestras independientes, la desviación estándar estimada de la diferencia entre las medias dependía de las desviaciones estándar de los pesos antes y después del programa. Como ambas desviaciones eran relativamente grandes, $\hat{\sigma}_{\bar{x}_1-\bar{x}_2}$ también era grande y, en consecuencia, 19.7 ya no es significativamente mayor que 17. La prueba de muestras por pares controla esta variabilidad inicial y final de los pesos al enfocar su atención solamente en los cambios individuales de peso. Debido a esto, pudo detectar mejor la significancia de la pérdida de peso.

¿Debemos tratar las muestras como dependientes o como independientes?

Concluimos esta sección con dos ejemplos que muestran cuándo tratar dos muestras de igual tamaño como dependientes o independientes:

1. Un departamento de fomento agrícola del gobierno de Estados Unidos desea determinar si una nueva semilla híbrida de maíz tiene una mayor producción que la vieja variedad estándar. Si el departamento pide a 10 granjeros que registren la producción de un acre sembrado con la nueva variedad de semilla y a otros 10 que registren la producción de un acre plantado con la vieja variedad, las dos muestras son independientes. Sin embargo, si le pide a 10 granjeros que planten un acre con cada variedad de semilla y registren los resultados, entonces las muestras

Tabla 9-5 Medias y varianzas antes y después

Muestra	Tamaño	Media	Varianza
Antes	10	202.5	253.61
Después	10	182.8	201.96

son dependientes, y la prueba de diferencia por pares es la adecuada. En el último caso, se pueden controlar las diferencias debidas al tipo de fertilizante e insecticida, la cantidad de lluvia y otros, debido a que cada granjero trata sus dos acres de la misma manera. En consecuencia, cualquier diferencia en la producción se puede atribuir exclusivamente a la variedad plantada.

2. La directora de recursos secretariales de una oficina grande de abogados desea determinar si la velocidad de captura de un documento depende del tipo de procesador de textos que utilice una secretaria. Si prueba a siete secretarias que utilizan el programa PicosoftWrite y siete que utilizan WritePerfect, trataría a sus muestras como independientes. Si prueba a las mismas siete secretarias dos veces (cada vez con un procesador distinto), entonces las dos muestras son dependientes. En la prueba de diferencias apareadas, las diferencias entre las secretarias no son un factor adicional, y las diferencias entre las velocidades de captura se puede atribuir a los diferentes procesadores de palabras.

SUGERENCIAS Y SUPOSICIONES

A menudo, al probar las diferencias entre medias tiene sentido tomar muestras que no sean independientes entre sí. Por ejemplo, si se tratara de medir el efecto de un antioxidante en las tuberías de metal, por lo general, se tomaría una muestra de la oxidación en las tuberías antes y después de aplicar el antioxidante. Al hacerlo se controlarían los efectos para distintos sitios, calor y humedad. Puesto que algunas tuberías se incluirían dos veces, las muestras no serían independientes. Sugerencia: si se mide la oxidación en cada tubería antes y seis meses después de la aplicación, se tiene una sola muestra de los gramos de oxidación que aparecieron desde la aplicación.

Ejercicios 9.4

Ejercicios de autoevaluación

EA 9-5 Sherri Welch es una ingeniera de control de calidad de la división de limpiaparabrisas de Emsco, Inc. La empresa estudia dos nuevos hules sintéticos para sus limpiadores y Sherri es la encargada de determinar si los hules con los dos nuevos compuestos se desgastan igual. Equipa 12 autos de empleados de Emsco con un limpiador de cada uno de los compuestos. En los autos 1 a 6, el limpiador derecho está fabricado con el compuesto A y el izquierdo con el B; en los autos 7 a 12, el compuesto A se colocó en el limpiador izquierdo. Los carros se usaron en condiciones normales de operación hasta que los hules no realizaban un trabajo satisfactorio al limpiar el parabrisas. Los datos presentados se refieren a la vida útil (en días) de los hules. Para $\alpha = 0.05$, ¿es igual el desgaste de los dos compuestos?

Auto	1	2	3	4	5	6	7	8	9	10	11	12
Limp. izq.	162	323	220	274	165	271	233	156	238	211	241	154
Limp. der.	183	347	247	269	189	257	224	178	263	199	263	148

EA 9-6 Se pidió a nueve distribuidores de componentes de computadora en un área metropolitana importante que proporcionaran sus precios de dos impresoras a color de inyección de tinta. Los resultados de la encuesta se dan en la tabla (con precios en dólares). Para $\alpha = 0.05$, ¿es razonable asegurar que en promedio la impresora Apson es menos costosa que la Okaydata?

Distribuidor	1	2	3	4	5	6	7	8	9
Precio de Apson	$250	319	285	260	305	295	289	309	275
Precio de Okaydata	$270	325	269	275	289	285	295	325	300

Aplicaciones

■ **9-14** Los datos de la tabla corresponden a una muestra aleatoria de nueve empresas tomadas de la sección "Digest of Earnings Reports" (Resumen de Informes de Ingresos) del *The Wall Street Journal* del 6 de febrero de 1992:

a) Encuentre el cambio medio en los ingresos por acción, entre 1991 y 1992.
b) Encuentre la desviación estándar del cambio y la desviación estándar del error de la media.
c) ¿Fueron diferentes los ingresos medios por acción en 1991 y 1992? Pruebe con un nivel $\alpha = 0.02$.

Empresa	1	2	3	4	5	6	7	8	9
Ingreso de 1991	1.38	1.26	3.64	3.50	2.47	3.21	1.05	1.98	2.72
Ingreso de 1992	2.48	1.50	4.59	3.06	2.11	2.80	1.59	0.92	0.47

■ **9-15** Jeff Richardson, el encargado de recepción de un distribuidor de productos químicos, se enfrenta con el problema continuo de recibir tubos de ensaye, platos Petri y matraces rotos. Jeff determinó algunas precauciones adicionales de empaque que se pueden tomar para prevenir la rotura de las piezas y ha pedido al director de adquisiciones que informe a los proveedores de las nuevas medidas. En la tabla se dan los datos de 8 proveedores en términos del número promedio de piezas rotas por envío. ¿Indican los datos, para $\alpha = 0.05$, que las nuevas medidas han disminuido el número promedio de piezas rotas?

Proveedor	1	2	3	4	5	6	7	8
Antes	16	12	18	7	14	19	6	17
Después	14	13	12	6	9	15	8	15

■ **9-16** Additives-R-Us desarrolló un aditivo para mejorar la eficiencia del combustible en camiones de carga pesada. Probaron el aditivo seleccionando al azar 18 camiones y agrupándolos en nueve pares. En cada par, ambos camiones llevaban el mismo tipo de carga en la misma carretera, pero sólo se puso el nuevo aditivo a uno de ellos. Cada par siguió rutas distintas y llevó diferentes cargas. ¿Indican los datos, al nivel $\alpha = 0.01$, que los camiones que usaron aditivo lograron una eficiencia en el uso de combustible significativamente mejor que los camiones con combustible normal?

Par	1	2	3	4	5	6	7	8	9
Normal	5.7	6.1	5.9	6.2	6.4	5.1	5.9	6.0	5.5
Aditivo	6.0	6.2	5.8	6.6	6.7	5.3	5.7	6.1	5.9

■ **9-17** El club deportivo Aquarius Health anuncia un riguroso programa de acondicionamiento físico. El club asegura que después de un mes de seguir el programa, un participante promedio será capaz de hacer 8 "lagartijas" más en 2 minutos que las que podía hacer al principio. ¿La muestra aleatoria de 10 participantes en el programa, cuyos datos se dan en la tabla siguiente, apoya la afirmación del club? Utilice un nivel de significancia de 0.025.

Participante	1	2	3	4	5	6	7	8	9	10
Antes	38	11	34	25	17	38	12	27	32	29
Después	45	24	41	39	30	44	30	39	40	41

■ **9-18** Donna Rose es supervisora de producción de la línea de ensamble de unidades de disco de Winchester Technologies. Recientemente, Winchester instaló un sistema de audio para música ambiental en sus instalaciones, con la idea de que la música relajara a sus obreros y condujera a una mayor productividad. Donna duda de esta hipótesis, teme que la música sea un foco de distracción y produzca una baja en la productividad. Muestreó la producción semanal de los mismos seis trabajadores antes de tener música ambiental y después instalar el sistema. Sus datos se presentan a continuación. A un nivel $\alpha = 0.02$, ¿ha cambiado la producción promedio?

Empleado	1	2	3	4	5	6
Semana sin música	219	205	226	198	209	216
Semana con música	235	186	240	203	221	205

■ **9-19** La velocidad de transmisión de un módem se mide en baudios, que se definen como el número de bits por segundo que puede transmitir. Debido a la intervención de varios factores técnicos, la rapidez de transmisión real varía de un archivo a otro. Anne Evans está en proceso de adquirir un módem de 28,800 baudios. Al probar dos de ellos para decidir cuál comprar, transmitió 7 archivos elegidos al azar utilizando ambos módems y registró las siguientes velocidades de transmisión (en miles de baudios).

Archivo	1	2	3	4	5	6	7
Haynes Ultima 28.8	9.52	10.17	10.33	10.02	10.72	9.62	9.17
Extel PerFAXtion 28.8	10.92	11.46	11.18	12.21	10.42	11.36	10.47

La revista *PC Reports* afirma que en pruebas hechas por su equipo se ha encontrado que el Extel PerFAXtion es significativamente más rápido que el Haynes Ultima. Para $\alpha = 0.01$, ¿los resultados obtenidos por Anne confirman la conclusión de la revista?

Soluciones a los ejercicios de autoevaluación

EA 9-5

Auto	1	2	3	4	5	6	7	8	9	10	11	12
Limp. A	183	347	247	269	189	257	233	156	238	211	241	154
Limp. B	162	323	220	274	165	271	224	178	263	199	263	148
Diferencia	21	24	27	−5	24	−14	9	−22	−25	12	−22	6

$$\bar{x} = \frac{\Sigma x}{n} = \frac{35}{12} = 2.9167 \text{ días}$$

$$s^2 = \frac{1}{n-1}(\Sigma x^2 - n\bar{x}^2) = \frac{1}{11}(4397 - 12(2.9167)^2) = 390.45, \; s = \sqrt{s^2}$$

$$= 19.76 \text{ días}$$

$$\hat{\sigma}_{\bar{x}} = s/\sqrt{n} = 19.76/\sqrt{12} = 4.7042 \text{ días}$$

H_0: $\mu_A = \mu_B$ H_1: $\mu_A \neq \mu_B$ $\alpha = 0.05$

Los límites de la región de aceptación son $t = \pm 2.201$, o

$$\bar{x} = 0 \pm t\hat{\sigma}_{\bar{x}} = \pm 2.201(5.7042) = \pm 12.55 \text{ días}$$

Como el valor observado $t = \dfrac{\bar{x} - \mu_{H_0}}{\hat{\sigma}_{\bar{x}}} = \dfrac{2.9167 - 0}{5.7042} = 0.511 < 2.201$

(o $\bar{x} = 2.9167 < 12.55$), no ser rechaza H_0. Los dos compuestos no son significativamente diferentes respecto a la vida útil.

EA 9-6

Distribuidor	1	2	3	4	5	6	7	8	9
Precio de Apson	250	319	285	260	305	295	289	309	275
Precio de Okaydata	270	325	269	275	289	285	295	325	300
Diferencia	20	6	−16	15	−16	−10	6	16	25

$$\bar{x} = \frac{\Sigma x}{n} = \frac{46}{9} = \$5.1111$$

$$s^2 = \frac{1}{n-1}(\Sigma x^2 - n\bar{x}^2) = \frac{1}{8}(2{,}190 - 9(5.1111)^2) = 244.36, \; s = \sqrt{s^2} = \$15.63$$

$$\hat{\sigma}_{\bar{x}} = s/\sqrt{n} = 15.63/\sqrt{9} = \$5.21$$

H_0: $\mu_O = \mu_A$ H_1: $\mu_O = \mu_A$ $\alpha = 0.05$

El límite superior de la región de aceptación es $t = 1.860$, o

$$\bar{x} = 0 + t\hat{\sigma}_{\bar{x}} = 1.860(5.21) = \$9.69$$

Como el valor observado $t = \dfrac{\bar{x} - \mu_{H_0}}{\hat{\sigma}_{\bar{x}}} = \dfrac{5.1111 - 0}{5.21} = 0.981 < 1.860$ (o $\bar{x} = \$5.11 < \$9.69$), no se rechaza H_0. En promedio, la impresora de inyección de tinta Apson no es significativamente menos costosa que la Okaydata.

9.5 Pruebas para diferencias entre proporciones: muestras grandes

Suponga que está interesado en averiguar si el Partido Republicano de Estados Unidos es más fuerte en Nueva York que en California. O tal vez desearía saber si las mujeres tienen la misma posibili-

dad que los hombres de adquirir automóviles deportivos. Para llegar a conclusiones en situaciones como éstas, usted puede tomar muestras de cada uno de los dos grupos en cuestión (votantes en Nueva York y en California, o mujeres y hombres) y utilizar las proporciones de muestra para probar la diferencia entre las dos poblaciones.

El procedimiento general a seguir es muy parecido a lo que hicimos en la sección 9.2, cuando comparamos dos medias utilizando muestras independientes: estandarizamos la diferencia entre las dos proporciones de muestra y basamos nuestras pruebas en la distribución normal. La única diferencia importante se dará en la forma en que encontremos una estimación para el error estándar de la diferencia entre las dos proporciones de muestra. Veamos algunos ejemplos.

Pruebas de dos colas para diferencias entre proporciones

Considere el caso de una compañía que fabrica productos medicinales y que está probando dos nuevos compuestos destinados a reducir los niveles de presión sanguínea. Los compuestos se administran a dos conjuntos de animales de laboratorio. En el grupo uno, 71 de 100 animales respondieron a la droga 1 con niveles menores de presión arterial. En el grupo dos, 58 de 90 animales respondieron a la droga 2 con menores niveles de presión sanguínea. La compañía desea probar a un nivel de significancia de 0.05 si existe una diferencia en la eficacia de las dos medicinas. ¿De qué manera debemos proceder con respecto a este problema?

$\bar{p}_1 = 0.71$ ← Proporción muestral de éxitos con la droga 1
$\bar{q}_1 = 0.29$ ← Proporción muestral de fracasos con la droga 1
$n_1 = 100$ ← Tamaño de la muestra para probar la droga 1
$\bar{p}_2 = 0.644$ ← Proporción muestral de éxitos con la droga 2
$\bar{q}_2 = 0.356$ ← Proporción muestral de fracasos con la droga 2
$n_2 = 90$ ← Tamaño de la muestra para probar la droga 2

Paso 1: Establezca sus hipótesis, tipo de prueba y nivel de significancia

$H_0: p_1 = p_2$ ← Hipótesis nula: no existe diferencia entre las dos drogas
$H_1: p_1 \neq p_2$ ← Hipótesis alternativa: sí existe diferencia entre ellas
$\alpha = 0.05$ ← Nivel de significancia para probar esta hipótesis

Paso 2: Escoja la distribución apropiada y encuentre el valor crítico

La figura 9-8 ilustra esta prueba de hipótesis. Debido a que la administración de la compañía farmacéutica desea saber si existe una diferencia entre los dos compuestos, se trata de una prueba de dos colas. El nivel de significancia de 0.05 corresponde a las regiones sombreadas de la figura. Ambas muestras son suficientemente grandes para justificar el uso de la distribución normal para aproximar a la binomial. En la tabla 1 del apéndice podemos determinar que el valor crítico de z para 0.475 del área bajo la curva es 1.96.

Como en los ejemplos anteriores, empezamos por calcular la desviación estándar de la distribución de muestreo para la prueba de hipótesis. En este ejemplo, la distribución binomial es la distribución de muestreo correcta.

Paso 3: Calcule el error estándar y estandarice el estadístico de la muestra

Deseamos encontrar el *error estándar de la diferencia entre dos proporciones*; por tanto, debemos recordar la fórmula para el *error estándar de la proporción*:

$$\sigma_{\bar{p}} = \sqrt{\frac{pq}{n}} \qquad [7\text{-}4]$$

Utilizando esta fórmula y procediendo como lo hicimos en la ecuación 9-1 para el error estándar de la diferencia entre *dos medias*, obtenemos:

Error estándar de la diferencia entre dos proporciones

$$\sigma_{\bar{p}_1 - \bar{p}_2} = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}} \qquad [9\text{-}5]$$

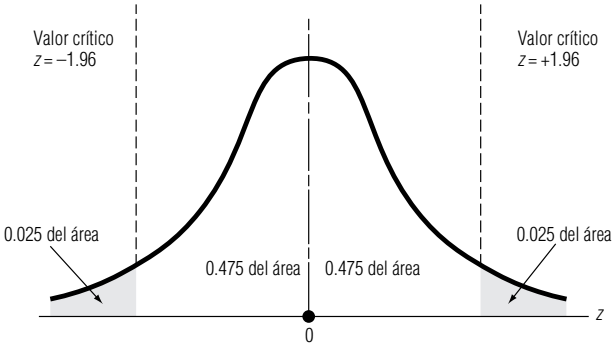


FIGURA 9-8

Prueba de hipótesis de dos colas de la diferencia entre dos proporciones al nivel de significancia de 0.05

Cómo estimar este error estándar

Para probar los dos compuestos, no conocemos los parámetros de la población p_1, p_2, q_1 y q_2, entonces, necesitamos estimarlos a partir de los estadísticos de la muestra $\bar{p}_1$, $\bar{p}_2$, $\bar{q}_2$ y $\bar{q}_2$. En este caso, podríamos suponer que la fórmula más práctica sería:

Error estándar estimado de la diferencia entre dos proporciones

Proporciones muestrales para la muestra 1 — Proporciones muestrales para la muestra 2

$$\hat{\sigma}_{\bar{p}_1 - \bar{p}_2} = \sqrt{\frac{\bar{p}_1\bar{q}_1}{n_1} + \frac{\bar{p}_2\bar{q}_2}{n_2}} \qquad [9\text{-}6]$$

Pero pensemos en esto un poco más. Después de todo, si establecemos la hipótesis de que *no hay diferencia* entre las dos proporciones de población, entonces la mejor estimación de la proporción global de éxitos en la población es, tal vez, la *proporción combinada de éxitos* en ambas muestras, esto es:

$$\text{Mejor estimación de la proporción global de éxitos en la población con la hipótesis de que las dos proporciones son iguales} = \frac{\text{número de éxitos en la muestra 1} + \text{número de éxitos en la muestra 2}}{\text{tamaño total de ambas muestras}}$$

Y en el caso de los dos compuestos, utilizamos esta ecuación:

Proporción global estimada de éxitos en dos poblaciones

$$\hat{p} = \frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2} \qquad [9\text{-}7]$$

$$= \frac{(100)(0.71) + (90)(0.644)}{100 + 90}$$

$$= \frac{71 + 58}{190}$$

$$= 0.6789 \leftarrow$$ Estimación de la proporción global de éxitos en las poblaciones combinadas utilizando las proporciones combinadas de ambas muestras ($\hat{q}$ sería $1 - 0.6789 = 0.3211$)

Ahora podemos modificar la ecuación 9-6 usando los valores de $\hat{p}$ y $\hat{q}$ tomados de la ecuación 9-7.

Error estándar estimado de la diferencia entre dos proporciones usando estimaciones combinadas de ambas muestras

Estimaciones de las proporciones de población usando proporciones combinadas de ambas muestras

$$\hat{\sigma}_{\bar{p}_1 - \bar{p}_2} = \sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}} \quad [9\text{-}8]$$

$$= \sqrt{\frac{(0.6789)(0.3211)}{100} + \frac{(0.6789)(0.3211)}{90}}$$

$$= \sqrt{\frac{0.2180}{100} + \frac{0.2180}{90}}$$

$$= \sqrt{0.004602}$$

$$= 0.0678$$

← Error estándar estimado de la diferencia entre dos proporciones

Estandarizamos la diferencia entre las dos proporciones de la muestra observadas, $\bar{p}_1 - \bar{p}_2$, dividiéndola entre el error estándar estimado de la diferencia entre dos proporciones:

$$z = \frac{(\bar{p}_1 - \bar{p}_2) - (p_1 - p_2)_{H_0}}{\hat{\sigma}_{\bar{p}_1 - \bar{p}_2}}$$

$$= \frac{(071 - 0.644) - 0}{0.0678}$$

$$= 0.973$$

Paso 4: Bosqueje la distribución y señale el valor de la muestra y los valores críticos

Paso 5: Interprete el resultado

Trazamos el valor estandarizado en una gráfica de la distribución de muestreo, como la figura 9-9.

En la figura 9-9 podemos ver que la diferencia estandarizada entre las dos proporciones de la muestra se encuentra dentro de la región de aceptación. Así, aceptamos la hipótesis nula y concluimos que las dos drogas nuevas producen efectos en la presión sanguínea que *no* son significativamente diferentes.

Pruebas de una cola para diferencias entre proporciones

Conceptualmente, la prueba de una cola para la diferencia entre dos proporciones de la población es parecida a la prueba de una cola para la diferencia entre dos medias. Suponga que con fines de im-

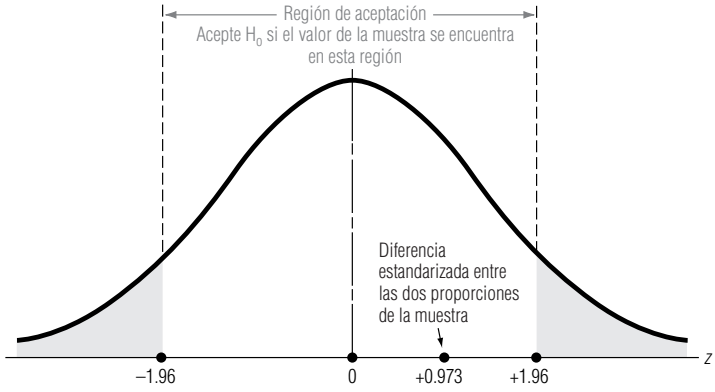


FIGURA 9-9

Prueba de hipótesis de dos colas de la diferencia entre dos proporciones al nivel 0.05 de significancia; se indican la región de aceptación y la diferencia estandarizada entre las proporciones de la muestra

puestos, el gobierno de una ciudad ha utilizado dos métodos para registrar las propiedades. El primero requiere que el dueño de la propiedad se presente para proporcionar la información; el segundo método permite que el dueño de la propiedad envíe por correo una forma fiscal con la información requerida. El alcalde de la ciudad piensa que el método que requiere la presencia de la persona produce menos errores que el otro. Autoriza un estudio de 50 listas hechas con el primer método y de 75 listas tomadas de los datos enviados por correo. El 10% de las formas llenadas ante el representante fiscal tiene errores; el 13.3% de las formas enviadas por correo tiene errores. Los resultados del muestreo pueden resumirse de la siguiente manera:

$\bar{p}_1 = 0.10$ ← Proporción de formas entregadas en persona que contienen errores
$\bar{q}_1 = 0.90$ ← Proporción de formas entregadas en persona que no contienen errores
$n_1 = 50$ ← Tamaño de la muestra de formas entregadas en persona
$\bar{p}_2 = 0.133$ ← Proporción de formas enviadas por correo que contienen errores
$\bar{q}_2 = 0.867$ ← Proporción de formas enviadas por correo que no contienen errores
$n_2 = 75$ ← Tamaño de la muestra de formas enviadas por correo

El alcalde de la ciudad desea probar, al nivel de significancia de 0.15, la hipótesis de que el método de entrega personal produce una proporción menor de errores. ¿Que deberá hacer?

Paso 1: Establezca sus hipótesis, tipo de prueba y nivel de significancia

$H_0: p_1 = p_2$ ← Hipótesis nula: no existe diferencia entre los dos métodos
$H_1: p_1 < p_2$ ← Hipótesis alternativa: el método de entrega personal tiene una proporción menor de errores que el método de entrega por correo

$\alpha = 0.15$ ← Nivel de significancia para probar la hipótesis

Paso 2: Escoja la distribución apropiada y encuentre el valor crítico

Con muestras de este tamaño, podemos utilizar la distribución normal estándar y la tabla 1 del apéndice para determinar el valor crítico de *z* para 0.35 del área bajo la curva (0.50 − 0.15). Este valor, 1.04, es el límite de la región de aceptación.

La figura 9-10 ilustra esta prueba de hipótesis. Como el alcalde de la ciudad desea probar si el método de entrega personal es mejor que el de envío por correo, lo apropiado es una prueba de una cola. Específicamente, se trata de una prueba *de cola izquierda*, ya que para rechazar la hipótesis nula, el resultado de la prueba debe caer en la región sombreada de la cola izquierda, lo cual indica que existen *significativamente menos errores* en las formas entregadas personalmente. Esta región sombreada de la figura 9-10 corresponde al nivel de significancia de 0.15.

Paso 3: Calcule el error estándar y estandarice la estadística de la muestra

Para estimar el *error estándar de la diferencia entre dos proporciones*, primero utilizamos las proporciones combinadas de ambas muestras para estimar la proporción global de éxitos:

$$\hat{p} = \frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2} \qquad [9\text{-}7]$$

$$= \frac{(50)(0.10) + (75)(0.133)}{50 + 75}$$

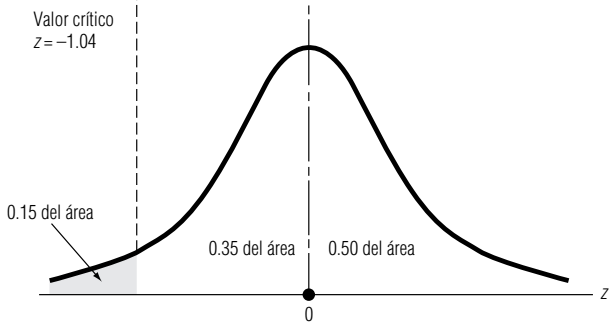


FIGURA 9-10

Prueba de hipótesis de una cola para la diferencia entre dos proporciones al nivel de significancia de 0.15

$$= \frac{5 + 10}{125}$$

$$= 0.12 \leftarrow \text{Estimación de la proporción global de éxitos en la población, utilizando proporciones combinadas de ambas muestras}$$

Esta respuesta se puede usar para calcular el error estándar de la diferencia entre las dos proporciones. Con la ecuación 9-8, obtenemos:

$$\hat{\sigma}_{\bar{p}_1 - \bar{p}_2} = \sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}} \qquad [9\text{-}8]$$

$$= \sqrt{\frac{(0.12)(0.88)}{50} + \frac{(0.12)(0.88)}{75}}$$

$$= \sqrt{\frac{0.10560}{50} + \frac{0.10560}{75}}$$

$$= \sqrt{0.00352}$$

$$= 0.0593 \leftarrow \text{Error estándar estimado de la diferencia entre dos proporciones, utilizando estimaciones combinadas}$$

Utilizamos el error estándar estimado de la diferencia, $\hat{\sigma}_{\bar{p}_1 - \bar{p}_2}$, para convertir la diferencia observada entre las dos proporciones de muestra, $\bar{p}_1 - \bar{p}_2$, en un valor estandarizado:

$$z = \frac{(\bar{p}_1 - \bar{p}_2) - (p_1 - p_2)_{H_0}}{\hat{\sigma}_{\bar{p}_1 - \bar{p}_2}}$$

$$= \frac{(0.10 - 0.133) - 0}{0.0593}$$

$$= -0.556$$

Paso 4: Bosqueje la distribución y señale el valor de la muestra y el valor crítico

Paso 5: Interprete el resultado

En la figura 9-11 se muestra el lugar donde se encuentra esta diferencia estandarizada, en comparación con el valor crítico.

Esta figura ilustra que la diferencia estandarizada entre las proporciones de la muestra se encuentra muy adentro de la región de aceptación y el alcalde de la ciudad debe aceptar la hipótesis nula de que no existe diferencia entre los dos métodos de recabar información. Así, si el método de envío por correo es considerablemente menos caro que el otro, el alcalde deberá pensar en aumentar su uso.

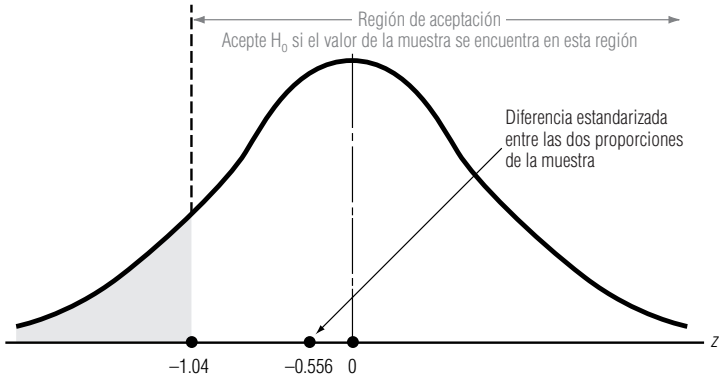


FIGURA 9-11

Prueba de hipótesis de una cola de la diferencia entre dos proporciones, al nivel de significancia de 0.15; se indica la región de aceptación y la diferencia estandarizada entre las proporciones de muestra

SUGERENCIAS Y SUPOSICIONES El procedimiento es casi igual que el usado al comparar las diferencias entre dos medias de muestras independientes. La única diferencia en este caso es que primero se usan las proporciones combinadas de ambas muestras para estimar la proporción global, después la respuesta sirve para estimar el error estándar de la diferencia entre las dos proporciones. Sugerencia: si la prueba se refiere a si una proporción es significativamente *diferente* de la otra, se usa la prueba de dos colas; si la prueba pregunta si una proporción es significativamente *mayor* o *menor* que la otra, entonces una prueba de una cola es adecuada.

Ejercicios 9.5

Ejercicios de autoevaluación

EA 9-7 Una cadena de hoteles grande intenta decidir si debe convertir más de sus cuartos en habitaciones de no fumar. El año pasado, en una muestra aleatoria de 400 huéspedes, 166 pidieron cuartos de no fumar. Este año, 205 huéspedes en una muestra de 380 prefirieron que no se fumara en su habitación. ¿Recomendaría que la cadena hotelera convirtiera más habitaciones en lugares de no fumar? Apoye su recomendación con la hipótesis adecuada a un nivel de significancia de 0.01.

EA 9-8 Dos áreas diferentes de una gran ciudad del este de Estados Unidos se estudian como sitios para abrir guarderías. De 200 casas encuestadas en una sección, la proporción de ellas en las que la madre trabaja tiempo completo es 0.52. En otra sección, el 40% de las 150 casas del estudio tiene madres que trabajan tiempo completo. Para un nivel de significancia de 0.04, ¿existe una diferencia significativa en las proporciones de madres que trabajan tiempo completo en las dos áreas de la ciudad?

Aplicaciones

■ **9-20** El viernes, aumentó el precio (avanzaron) de 11 acciones de una muestra aleatoria de 40 tomada de las 2,500 acciones negociadas en la Bolsa de Valores de Nueva York. En una muestra tomada el jueves, de 60 acciones de la misma Bolsa, 24 acciones avanzaron. A un nivel $\alpha = 0.10$, ¿puede llegar a la conclusión de que una proporción menor de las acciones de la Bolsa de Valores avanzaron el viernes con respecto al jueves?

■ **9-21** MacroSwift acaba de liberar al mercado un nuevo procesador de textos y la compañía está interesada en determinar si las personas en el grupo de edad 30-39 califican al programa de manera distinta a las del grupo 40-49. MacroSwift muestreó al azar a 175 personas del grupo 30-39 que compraron el producto y encontró que 87 calificaron al programa como excelente; de ellos, 52 comprarían una actualización. También muestreó a 220 personas del grupo 40-49 y encontró que 94 calificaron al software como excelente; de ellos, 37 comprarían una actualización. ¿Hay una diferencia significativa en las proporciones de personas en los dos grupos de edad que califican al programa como excelente al nivel $\alpha = 0.05$? ¿Es cierto el mismo resultado en cuanto a las proporciones de personas que planean comprar una actualización?

■ **9-22** Una planta de energía impulsada por carbón está considerando dos sistemas diferentes para reducir la emisión de contaminantes. El primer sistema reduce la emisión a niveles aceptables el 68% del tiempo, según 200 muestras de aire. El segundo sistema, más costoso, la reduce a niveles aceptables el 76% del tiempo, de acuerdo con 250 muestras. Si el sistema costoso es significativamente más efectivo que el otro al reducir la emisión de contaminantes a niveles aceptables, entonces la administración de la planta instalará el sistema costoso. ¿Qué sistema se debe instalar si la administración usa un nivel de significancia de 0.02 al tomar su decisión?

■ **9-23** Un grupo de investigadores médicos está llevando a cabo pruebas en pacientes para determinar la efectividad de una nueva medicina contra la hipertensión. Se eligieron al azar pacientes con alta presión sanguínea y se asignaron, también en forma aleatoria, a un grupo de control (que recibió un conocido medicamento contra la hipertensión) o al grupo de tratamiento (que recibió la nueva medicina). Los médicos registraron el porcentaje de pacientes cuya presión arterial se redujo a un nivel normal después de un año de tratamiento. Al nivel de significancia de 0.01, pruebe las hipótesis apropiadas para determinar si la nueva medicina es significativamente más efectiva para reducir la presión sanguínea que la medicina conocida.

Grupo	Proporción que mejoró	Número de pacientes
Tratamiento	0.45	120
Control	0.36	150

■ **9-24** La librería de la universidad se enfrenta a una competencia significativa con librerías fuera de ella y está considerando dirigir sus ventas a una generación específica para retener a los estudiantes como clientes. Se realizó un muestreo aleatorio de 150 estudiantes de primer año y 175 de segundo. Encontraron que el 46% de primer año y el 40% de segundo compraban todos sus libros de texto en la librería universitaria. Para $\alpha = 0.10$, ¿existe una diferencia significativa en las proporciones de estudiantes de primero y segundo año que compran todo en la librería de la universidad?

■ **9-25** Como parte de la preparación para las negociaciones de renovación del contrato colectivo de trabajo, el Sindicato Unido de Manufactureros hizo una investigación entre sus afiliados para ver si hay preferencia por un aumento grande en los beneficios para el retiro y un menor incremento al salario. En un grupo de 1,000 miembros masculinos que fueron entrevistados, 743 estaban a favor de un aumento en los beneficios de retiro. De 500 miembros femeninos del sindicato, 405 estaban a favor del aumento en los beneficios de retiro.

a) Calcule $\hat{p}$.
b) Calcule el error estándar de la diferencia entre las dos proporciones.
c) Pruebe la hipótesis de que igual proporción de hombres que de mujeres están a favor de un aumento en los beneficios de retiro. Utilice un nivel de significancia de 0.05.

Soluciones a los ejercicios de autoevaluación

EA 9-7 $n_1 = 400 \quad \bar{p}_1 = 0.415 \quad n_2 = 380 \quad \bar{p}_2 = 0.5395$

$H_0: p_1 = p_2 \quad H_1: p_1 < p_2 \quad \alpha = 0.01$

$$\hat{p} = \frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2} = \frac{400(0.415) + 380(0.5395)}{400 + 380} = 0.4757$$

$$\hat{\sigma}_{\bar{p}_1 - \bar{p}_2} = \sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sqrt{0.4757(0.5243)\left(\frac{1}{400} + \frac{1}{380}\right)} = 0.0358$$

El límite inferior de la región de aceptación es $z = -2.33$, o

$$\bar{p}_1 - \bar{p}_2 = 0 - z\hat{\sigma}_{\bar{p}_1 - \bar{p}_2} = -2.33(0.0358) = -0.0834$$

Como el valor observado $z = \dfrac{\bar{p}_1 - \bar{p}_2}{\hat{\sigma}_{\bar{p}_1 - \bar{p}_2}} = \dfrac{0.415 - 0.5395}{0.0358} = -3.48 < -2.33$ (o $\bar{p}_1 - \bar{p}_2 = -0.1245 < -0.0834$), se rechaza H_0. La cadena de hoteles debe convertir más cuartos en habitaciones de no fumar puesto que hay un incremento significativo en la proporción de huéspedes que piden estos cuartos respecto al año pasado.

EA 9-8 $n_1 = 200 \quad \bar{p}_1 = 0.52 \quad n_2 = 150 \quad \bar{p}_2 = 0.40$

$H_0: p_1 = p_2 \quad H_1: p_1 \neq p_2 \quad \alpha = 0.04$

$$\hat{p} = \frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2} = \frac{200(0.52) + 150(0.40)}{200 + 150} = 0.4686$$

$$\hat{\sigma}_{\bar{p}_1 - \bar{p}_2} = \sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sqrt{0.4686(0.5314)\left(\frac{1}{200} + \frac{1}{150}\right)} = 0.539$$

Los límites de la región de aceptación son $z = \pm 2.05$, o

$$\bar{p}_1 - \bar{p}_2 = 0 - z\hat{\sigma}_{\bar{p}_1 - \bar{p}_2} = \pm 2.05(0.0539) = \pm 0.1105$$

Como el valor observado $z = \dfrac{\bar{p}_1 - \bar{p}_2}{\hat{\sigma}_{\bar{p}_1 - \bar{p}_2}} = \dfrac{0.52 - 0.40}{0.0539} = 2.23 > 2.05$ (o $\bar{p}_1 - \bar{p}_2 = 0.12 > 0.1105$), se rechaza H_0. Las proporciones de madres que trabajan en las dos áreas difieren de manera significativa.

9.6 Valor P: otra manera de ver las prueba de hipótesis

En el trabajo realizado sobre pruebas de hipótesis hasta ahora, una de las primeras cosas que tuvimos que hacer fue escoger un nivel de significancia, α, para la prueba. Se ha vuelto tradicional seleccionar un nivel de significancia de $\alpha = 10$, 5, 2 o 1%, y casi todos nuestros ejemplos se han resuelto con esos niveles. Pero, ¿por qué utilizar solamente esos valores?

¿Cómo escogemos un nivel de significancia?

Cuando analizamos los errores tipo I y tipo II, vimos que la elección del nivel de significancia dependía de un trueque entre los costos de cada tipo de error. Si el costo de un error tipo I (rechazo incorrecto de H_0) es relativamente alto, intentamos evitar este tipo de error y, en consecuencia, escogemos un valor pequeño de α. Por otro lado, si un error tipo II (aceptación incorrecta de H_0) es relativamente más caro, es preferible cometer un error tipo I, y escogemos un valor alto de α. **Sin embargo, entender la naturaleza del trueque no nos dice cómo escoger un nivel de significancia.**

Decisiones antes de tomar una muestra

Cuando probamos las hipótesis:

$$H_0: \mu = \mu_{H_0}$$
$$H_1: \mu \neq \mu_{H_0}$$
$$\alpha = 0.05$$

tomamos una muestra, calculamos $\bar{x}$ y rechazamos H_0 si $\bar{x}$ está tan alejada de μ_{H_0} que la probabilidad de encontrarse un valor de $\bar{x}$ tan lejos (o más) de μ_{H_0} es menor que 0.05. En otras palabras, **antes de tomar la muestra** especificamos qué tan improbables deberán ser los resultados observados para que rechacemos H_0. Existe otra forma de enfocar la decisión de aceptar o rechazar H_0 que no requiere especificar el nivel de significancia antes de tomar la muestra. Veamos cómo es que funciona.

Valores de prueba P

Suponga que tomamos una muestra, calculamos $\bar{x}$ y luego nos preguntamos: "suponiendo que H_0 es cierta, ¿cuál es la probabilidad de obtener un valor de $\bar{x}$ a esta distancia o mayor de μ_{H_0}?" Esta probabilidad se conoce como *valor P*. **Mientras que antes nos preguntábamos, "¿es menor que α la probabilidad de lo que observamos?", ahora nos preguntamos, "¿qué tan improbable es el resultado que observamos?". Una vez que se informa el valor P de la prueba, *en ese momento* el tomador de decisiones puede evaluar los factores relevantes y decidir si acepta o rechaza H_0, sin que lo limite un nivel de significancia especificado.**

Otra ventaja

Otro beneficio de utilizar valores P es que proporcionan más información. Si se sabe que se ha rechazado H_0 al nivel $\alpha = 0.05$, solamente sabe que $\bar{x}$ se encuentra *al menos* 1.96 errores estándar alejado de μ_{H_0}. Sin embargo, un valor P de 0.05 nos dice que $\bar{x}$ está *exactamente* a 1.96 errores estándar de μ_{H_0}. Veamos un ejemplo.

Valores P de dos colas cuando se conoce σ

Se utiliza una máquina para cortar bloques de queso suizo con un peso específico. Con base en una larga experiencia, se ha observado que el peso de los bloques sigue una distribución normal con desviación estándar de 0.3 onzas. La máquina está ajustada para cortar bloques de 12 onzas. En una muestra de nueve bloques se encontró que el peso promedio es 12.25 onzas. ¿Debemos concluir que la máquina cortadora necesita reajustarse?

Planteamiento del problema en símbolos

Escrito con símbolos, los datos del problema son:

$\mu_{H_0} = 12$ ← Valor hipotético de la media de la población
$\sigma = 0.3$ ← Desviación estándar de la población
$n = 9$ ← Tamaño de la muestra
$\bar{x} = 12.25$ ← Media de la muestra

Las hipótesis que deseamos probar son:

$H_0: \mu = 12$ ← Hipótesis nula: el peso medio verdadero de la población es 12 onzas
$H_1: \mu \neq 12$ ← Hipótesis alternativa: el peso medio verdadero de la población no es de 12 onzas

Como se trata de una prueba de dos colas, el valor P es la probabilidad de observar un valor de $\bar{x}$ alejado de la media 12 (en cualquier dirección) hasta 12.25, si H_0 es cierta. En otras palabras, el valor P es la probabilidad de obtener $\bar{x} \geq 12.25$ o $\bar{x} \leq 11.75$ si H_0 es cierta. Para encontrar esta probabilidad, primero utilizamos la ecuación 6-1 para calcular el error estándar de la media:

Cálculo del error estándar de la media

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \qquad [6\text{-}1]$$

$$= \frac{0.3}{\sqrt{9}}$$

$$= \frac{0.3}{3}$$

$$= 0.1 \text{ onza} \leftarrow \text{Error estándar de la media}$$

Éste se usa para convertir $\bar{x}$ en un resultado z estándar:

Búsqueda del resultado z del valor P

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} \qquad [6\text{-}2]$$

$$= \frac{12.25 - 12}{0.1}$$

$$= \frac{0.25}{0.1}$$

$$= 2.5$$

En la tabla 1 del apéndice observamos que la probabilidad de que z sea mayor que 2.5 es $0.5000 - 0.4938 = 0.0062$. En consecuencia, como se trata de una prueba de dos colas, el valor P es $2(0.0062) = 0.0124$. La figura 9-12 ilustra estos resultados. Dada esta información, el empacador de queso puede ahora decidir si recalibra la máquina (rechaza H_0) o no (acepta H_0).

Relación entre los valores P y los niveles de significancia

¿Cómo se relaciona esto con lo que hicimos cuando especificamos un nivel de significancia? Si hubiéramos adoptado un nivel de significancia de $\alpha = 0.05$, habríamos rechazado H_0. Esto puede verse en la figura 9-12. Para un nivel de significancia de $\alpha = 0.05$, rechazamos H_0 si $\bar{x}$ se encuentra tan alejada de μ_{H_0} menos de 0.05 del área bajo la curva se queda en las dos colas. Como el valor observado de $\bar{x} = 12.25$ sólo deja 0.0124 del área total en las colas, rechazaríamos H_0 a un nivel de significancia $\alpha = 0.05$. [También puede verificar este resultado observando en la tabla 1 del apéndice que los valores críticos z para $\alpha = 0.05$ son ± 1.96. Así, el valor estandarizado de $\bar{x}$ (2.5) *está fuera* de la región de aceptación.]

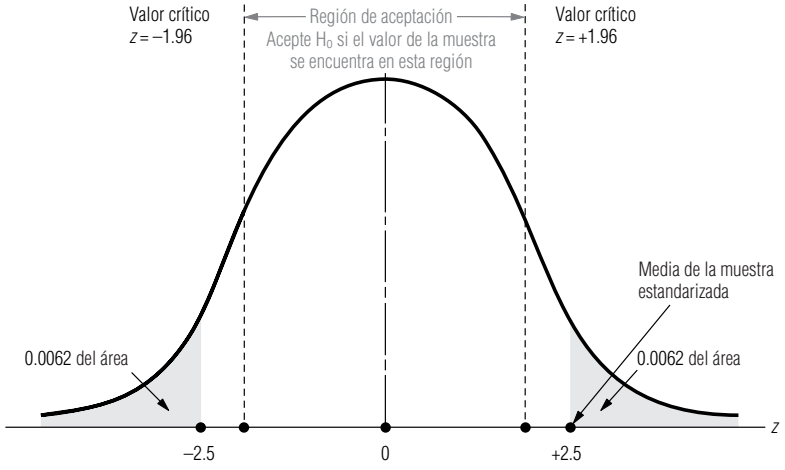


FIGURA 9-12

Prueba de hipótesis de dos colas; indica el valor P de 0.0124 (en ambas colas combinadas)

De manera parecida, podemos ver que a un nivel de significancia de $\alpha = 0.01$, hubiéramos aceptado H_0 debido a que $\bar{x} = 12.25$ deja más de 0.01 del área bajo la curva en las colas (en este caso, los valores críticos z para $\alpha = 0.01$ serían ± 2.58, y el valor estandarizado de $\bar{x}$, 2.5, se estaría *dentro* de la región de aceptación). De hecho, a cualquier nivel de α mayor que 0.0124 rechazaríamos la hipótesis nula H_0. **Así, vemos que el valor P es precisamente el nivel máximo de significancia para el cual aceptaríamos H_0.**

Valores P para otras condiciones

En nuestro ejemplo, hicimos una prueba de hipótesis de dos colas utilizando la distribución normal. ¿Cómo hubiéramos procedido con otras condiciones?

Valores P de una cola

1. Si σ fuera conocida y estuviéramos realizando una prueba de una cola, habríamos calculado el valor P exactamente de la misma manera, con la excepción de que no multiplicaríamos por dos la probabilidad obtenida en la tabla 1 del apéndice, pues esa tabla da las probabilidades de una cola directamente.

Uso de la distribución *t*

2. Si no se conociera σ, habríamos utilizado la distribución t con $n - 1$ grados de libertad y la tabla 2 del apéndice. Esta tabla da probabilidades de dos colas, pero sólo unas cuantas, de modo que no podemos obtener valores P exactos. Por ejemplo, para una prueba de dos colas, si $\mu_{H_0} = 50$, $\bar{x} = 49.2$, $s = 1.4$ y $n = 16$, encontramos que:

$$\hat{\sigma}_{\bar{x}} = \frac{\hat{\sigma}}{\sqrt{n}} \qquad [7\text{-}6]$$

$$= \frac{1.4}{\sqrt{16}}$$

$$= 0.35$$

y que $\bar{x}$ se encuentra a 2.286 errores estándar abajo de μ_{H_0} [$(49.2 - 50)/0.35 = -2.286$]. En el renglón de 15 grados de libertad en la tabla 2 del apéndice, vemos que 2.286 se encuentra entre 2.131 ($\alpha = 0.05$) y 2.602 ($\alpha = 0.02$). Nuestro valor P es entonces algún valor que está entre 0.02 y 0.05, pero no podemos ser más precisos.

Valores P en otros contextos

La mayoría de los paquetes estadísticos de computadora dan como resultado valores P exactos, no nada más para pruebas de medias basadas en la distribución normal, sino también para otras pruebas como ji-cuadrada y análisis de varianza (que estudiaremos en el capítulo 11) y pruebas en el contexto de la regresión lineal (que se verán en los capítulos 12 y 13). El análisis hecho en esta sección le permitirá entender los valores P también en esos contextos. Aunque se tengan diferentes estadísticos y distribuciones, las ideas son las mismas.

SUGERENCIAS Y SUPOSICIONES

Los valores P y las computadoras han suprimido la necesidad de buscar valores en las tablas de la distribución z o t, y eliminan el trabajo tedioso de las pruebas de hipótesis. Advertencia: cuanto *más pequeño* sea el valor P, mayor será la significancia del estudio. Sugerencia: se puede evitar la confusión aquí al recordar que un valor P es la probabilidad de que el resultado obtenido haya podido ocurrir por el error de muestreo; así, los valores P más pequeños significan menor posibilidad de error de muestreo y mayor significancia.

Ejercicios

Ejercicios de autoevaluación

EA 9-9 El Instituto del Café asegura que más del 40% de los adultos de Estados Unidos toma una taza de café en el desayuno. Una muestra aleatoria de 450 individuos reveló que 200 de ellos toman café de manera ha-

bitual en el desayuno. ¿Cuál es el valor P para una prueba de hipótesis que busca mostrar que la afirmación del Instituto del Café es correcta? (*Sugerencia:* pruebe H_0: $p = 0.4$, contra H_1: $p > 0.4$.)

EA 9-10 ¿Cuál es el valor P aproximado para la prueba del ejercicio de verificación 9-3?

Aplicaciones

- **9-26** Una vendedora de automóviles usados piensa que un fabricante de llantas exagera cuando afirma que sus llantas tienen una duración de 40,000 millas. Registra cuidadosamente el número de millas obtenido de una muestra de 64 llantas. Obtiene una media de 38,500 millas. El fabricante había calculado desviación estándar de la vida de todas las llantas de este tipo en 7,600 millas. Suponiendo que el número de millas tiene una distribución normal, determine el nivel de significancia más alto al cual aceptaríamos la afirmación del fabricante, es decir, el nivel al cual no concluiríamos que el número de millas es significativamente menor que 40,000 millas.
- **9-27** El Departamento de Transporte de Carolina del Norte asegura que, cuando mucho, el 18% de los autos excede 70 mph en la interestatal 40, entre Raleigh y Durham. Una muestra aleatoria de 300 automóviles encontró que 48 autos excedían las 70 mph. ¿Cuál es el valor P para una prueba de hipótesis que busca mostrar que la afirmación del DTCN es correcta?
- **9-28** En el taller mecánico de Kelly utilizan una sierra de motor para cortar el tubo que se usa en la manufactura de dispositivos de medición de presión. La longitud de los segmentos de tubo está distribuida normalmente con una desviación estándar de 0.06 pulg. Se cortaron 25 piezas de tubo con la sierra calibrada para secciones de 5.00 pulg. Cuando se midieron estas piezas, se encontró que su longitud media era 4.97 pulg. Utilice valores P para determinar si la máquina debe ser recalibrada debido a que la longitud media es significativamente diferente de 5.00 pulg.
- **9-29** Un centro de asesoría académica anuncia que el 80% de las veces, su curso de regularización incrementa el resultado individual de un alumno que presenta los exámenes de admisión a la universidad al menos 50 puntos, respecto al resultado combinado de las pruebas oral y escrita. Lisle Johns, gerente de comercialización del centro de asesoría, desea saber si su anuncio es razonable. Revisó los resultados de 125 estudiantes que tomaron el curso y encontró que 94 de ellos, efectivamente, aumentaron sus resultados 50 puntos o más. Utilice valores P para determinar si los anuncios del centro de asesoría deben cambiarse debido a que el porcentaje de estudiantes que aumentan su calificación es significativamente distinto del 80%.
- **9-30** ¿Cuál es el valor P para la prueba del ejercicio 9-2?
- **9-31** ¿Cuál es el valor P para la prueba del ejercicio 9-3?
- **9-32** ¿Cuál es el valor P aproximado para la prueba del ejercicio 9-8?
- **9-33** ¿Cuál es el valor P aproximado para la prueba del ejercicio 9-11 ?
- **9-34** ¿Cuál es el valor P aproximado para la prueba del ejercicio 9-14?
- **9-35** ¿Cuál es el valor P aproximado para la prueba del ejercicio 9-15?
- **9-36** ¿Cuál es el valor P para la prueba del ejercicio 9-22?
- **9-37** ¿Cuál es el valor P para la prueba del ejercicio 9-25?

Soluciones a los ejercicios de autoevaluación

EA 9-9 $n = 450 \qquad \bar{p} = 200/450 = 0.4444$

H_0: $p = 0.4 \qquad H_1$: $p > 0.4$

El valor P es la probabilidad de que $\bar{p} \geq 0.4444$, es decir,

$$P\left(z \geq \frac{0.4444 - 0.4}{\sqrt{0.4(0.6)/450}}\right) = P(z \geq 1.92) = 0.5 - 0.4726 = 0.0274$$

EA 9-10 De la solución del ejercicio EA 9-3, se tiene $t = -2.766$, con $12 + 9 - 2 = 19$ grados de libertad. Del renglón para 19 grados de libertad de la tabla 2 del apéndice, se ve que -2.766 está entre -2.861 (que corresponde a una probabilidad de $0.01/2 = 0.005$ en la cola inferior) y -2.539 (que corresponde a una probabilidad de $0.02/2 = 0.01$ en la cola inferior). Por tanto, el valor P para la prueba está entre 0.005 y 0.01.

9.7 Uso de computadoras para las pruebas de hipótesis

Cuando se diseñó el examen final del curso de estadística impartido durante el otoño de 1995, se esperaba que la calificación promedio aproximada fuera 75% (56.25 puntos de un resultado máximo posible de 75). Probemos (a un nivel $\alpha = 0.05$) si ocurrió lo que esperábamos.

Planteamiento del problema en símbolos

$$H_0: \mu = 56.25 \leftarrow \text{El examen tenía la dificultad deseada}$$
$$H_1: \mu \neq 56.25 \leftarrow \text{No se logró la dificultad deseada}$$
$$\alpha = 0.05 \leftarrow \text{Nivel de significancia para esta prueba}$$

Interpretación de los resultados

En la figura 9-13 utilizamos Minitab para analizar los datos del apéndice 10. El valor observado de t para esta prueba fue -15.45, con un valor P (de dos colas) asociado de 0.0000. Como este valor P es menor que nuestro nivel de significancia de $\alpha = 0.05$, debemos rechazar H_0 y concluir que la prueba no logró el nivel deseado de dificultad (de hecho, el examen resultó ser mucho más difícil de lo que intentamos).

La universidad ha estado recibiendo muchas quejas sobre la calidad de la enseñanza que imparten los estudiantes de posgrado que son ayudantes de profesor. Como resultado, nos preguntamos si los estudiantes que acudían a secciones que impartían los ayudantes salían más bajos en el examen que aquellos que iban a las secciones impartidas por profesores. Si las secciones de ayudantes son la muestra 1 y las de profesores la muestra 2, entonces las hipótesis apropiadas para probar esta preocupación son:

Establecimiento del problema de manera simbólica

$$H_0: \mu_1 = \mu_2 \leftarrow \text{La preocupación no está apoyada por los datos}$$
$$H_1: \mu_1 < \mu_2 \leftarrow \text{La preocupación está apoyada por los datos}$$

La salida de Minitab para este prueba se da en la figura 9-14. Observe que los resultados de la prueba se presentan suponiendo que las varianzas de las dos poblaciones son iguales. Si podemos suponer esto, entonces la prueba que reporta Minitab es la prueba estudiada en las páginas 376-377, donde se calculó una estimación combinada para σ^2.

FIGURA 9-13

Uso de Minitab para probar hipótesis sobre una media de población

```
Prueba T de la media

Prueba de mu = 56.250 vs mu no = 56.250

Variable     N     Media   DesvEst  EEst Media       T   Valor P
Final      199    45.281    10.014       0.710  -15.45    0.0000
```

FIGURA 9-14

Uso de Minitab para probar hipótesis sobre la diferencia entre dos medias

```
Prueba T de dos muestras e intervalo de confianza

Dosmuestras T para FINAL
NUMINSTR        N    Media   DesvEst    EEst Media
1              89    44.93      9.76          1.0
2             110     45.6      10.2         0.98

I.C. 95% para mu 1 - mu 2: (-3.5, 2.19)

Prueba T mu 1 = mu 2 (vs <): T= -0.44 P=0.33 GL= 197

Ambas usan DesvEst conjunta = 10.0
```

Interpretación de los resultados

¿Qué debe concluirse respecto a si los estudiantes de las secciones impartidas por ayudantes están o no en desventaja? El valor P obtenido para esta prueba de una cola (0.33) es bastante grande en relación con los niveles de significancia típicos que se han usado (0.10, 0.05, 0.01, etc.) de manera que no se puede rechazar H_0. Los datos no apoyan la preocupación expresada en las quejas recibidas por la universidad.

Los datos del apéndice 11 contienen los ingresos del último trimestre de 1988 para 224 compañías, además de los correspondientes al último trimestre de 1989. Debido a que la economía de Estados Unidos no había sufrido una recesión desde 1982, para 1989 muchos economistas esperaban que se desacelerara. Se puede verificar (para $\alpha = 0.10$) si el cambio de un año a otro en los ingresos en el último trimestre es evidencia de que sus expectativas eran correctas. En la figura 9-15, se usó Minitab para probar esto.

Establecimiento del problema en símbolos

$$H_0: \mu_{1989} = \mu_{1988} \leftarrow \text{La economía no se desaceleró}$$
$$H_1: \mu_{1989} < \mu_{1988} \leftarrow \text{La economía se desaceleró}$$
$$\alpha = 0.10 \leftarrow \text{Nivel de significancia para esta prueba}$$

Para realizar esta prueba de diferencia entre pares, primero se restan los ingresos del último trimestre de 1988 de los del último trimestre de 1989 y se guardan los resultados en una variable llamada "CAMBIO". Al observar CAMBIO, la hipótesis se convierte en

$$H_0: \mu = 0 \leftarrow \text{La economía no se desaceleró}$$
$$H_1: \mu < 0 \leftarrow \text{La economía se desaceleró}$$
$$\alpha = 0.10 \leftarrow \text{Nivel de significancia para esta prueba}$$

Interpretación de los resultados

En la figura 9-15 se observa que el valor P reportado para la prueba de una cola es 0.28. Como este valor P es mayor que el nivel de significancia $\alpha = 0.10$, no se puede rechazar H_0. Los ingresos del último trimestre no disminuyeron en forma significativa de un año a otro, de modo que no proporcionan evidencia de una economía desacelerada.

En esta prueba se usaron las 224 compañías. Sin embargo, quizá los cambios en los ingresos de las compañías relativamente grandes cuyas acciones están en la lista de la Bolsa de Valores de Nueva York (NYSE) sean diferentes de los cambios en los ingresos de compañías más pequeñas que ope-

FIGURA 9-15

Uso de Minitab para una prueba de diferencias por pares de medias

```
Prueba T de la media

Prueba de mu = 0.0000 contra mu < 0.0000

Variable     N     Media  DesvEst  EEst Media       T   Valor P
CAMBIO     224   -0.0354   0.8967      0.0599   -0.59      0.28
```

FIGURA 9-16

Uso de Minitab para una prueba de medias con muestras independientes

```
Prueba T de dos muestras e intervalo de confianza

Dosmuestras T para CAMBIO
NYONO     N     Media   DesvESt   EEst Media
0       149    -0.061     0.645        0.053
1        75      0.02      1.26         0.15

I.C. 95% para mu 0 - mu 1: ( -0.327, 0.17)

Prueba T mu 0 = mu 1 (contra no =): T= -0.60 P=0.55 GL= 222

Ambas usan DesvEst conjunta = 0.898
```

ran en la American Stock Exchange o de manera privada. Si difieren, tal vez 1 sea mejor indicador de la dirección de la economía. Para verificar esto, se pueden dividir las 224 compañías en dos grupos (NYSE y OTROS) y obtener muestras independientes para probar la diferencia en los valores de las medias de CAMBIO en los dos grupos. Probemos esto para $\alpha = 0.02$.

Establecimiento del problema en símbolos

$$H_0: \mu_{NYSE} = \mu_{OTRO} \leftarrow \text{Ambos cambios en las medias son iguales}$$
$$H_1: \mu_{NYSE} \neq \mu_{OTRO} \leftarrow \text{Ambos cambios en las medias son diferentes}$$
$$\alpha = 0.02 \leftarrow \text{Nivel de significancia para esta prueba}$$

En la figura 9-16 se usó Minitab para realizar esta prueba. La variable CAMBIO se convirtió de sus valores en caracteres O, A y N a los valores numéricos 0, 0 y 1 en una nueva variable llamada NYONO.

Interpretación de resultados

En este caso en particular el valor P de 0.55 para la prueba de la diferencia de los cambios en las medias de NYSE y OTRO es mucho mayor que el nivel de significancia $\alpha = 0.02$, de manera que no se puede rechazar H_0. No existe suficiente evidencia para garantizar más investigación del uso del cambio de un año a otro en los ingresos del último trimestre como evidencia de la dirección que tomará la economía.

Estadística en el trabajo

Loveland Computers

Caso 9: Prueba de hipótesis de dos muestras Cuando Lee Azko observó los resultados de la investigación telefónica efectuada por el departamento de mercadotecnia de Loveland Computers, tuvo la impresión de que algo no andaba bien.

"Hmmm, ya no tienes en la computadora los datos del 'gasto total en software', ¿verdad, Margot?", le preguntó Lee a la jefa del departamento.

"¡Claro que sí! Yo guardo *todo*", le respondió Margot. "Están en un archivo de hoja de cálculo en aquella computadora. Estuve encargada de procesar toda la información el verano pasado. ¿Por qué necesitas ver esos datos?"

"Bueno, dame un minuto y te lo diré", respondió Lee, al tiempo que encendía la computadora. Después de unos minutos de murmurar cosas mientras manipulaba el teclado, Lee se recargó en el respaldo de su asiento. "¡Como lo había pensado! Ven a ver esto. Parece como si en realidad tuviéramos dos tipos de clientes. ¿Ves cómo se forman dos picos distintos en esta gráfica?"

"Creo que debimos haber hecho más que imprimir la media y la desviación estándar el verano pasado", comentó Margot desanimada. "Me imagino que esto significa que los datos no son buenos."

"No necesariamente", dijo Lee más optimista. "Te apuesto a que tus 'grandes compradores' son los clientes ejecutivos de la compañía y el pico más pequeño es de los usuarios domésticos. No tendrías manera de saber de qué categoría viene la respuesta, ¿o sí?"

"Bueno, capturamos la información de manera automática", dijo Margot, inclinándose para ver mejor la gráfica en la pantalla. "Si te fijas en la primera columna, verás que corresponde al número de cliente. Todos los clientes ejecutivos tienen un número que empieza con un 1, y el número de los clientes domésticos empieza con un 2."

"Déjame copiar este archivo en un disquete", dijo Lee y abrió su portafolios. "Por la tarde te tendré la respuesta."

Preguntas de estudio: ¿Qué gráfica hizo Lee utilizando la hoja de cálculo? ¿Qué hipótesis está probando y cuál es la prueba estadística apropiada? ¿Se trata de un problema de una cola o de dos colas?

Ejercicio de base de datos computacional

HH Industries

Laurel tocó en la puerta entreabierta de la oficina de Hal. "Bienvenido", dijo ella a modo de saludo. "Supe que tuviste un buen viaje."

Hal levantó la mirada y le hizo señas para que entrara y se sentara. "Pues sí, fue magnífico. ¡Pero estoy exhausto! Nos pasamos la semana completa de pie."

Hal, Stan y John Raymond, el nuevo gerente de publicidad, habían regresado el día anterior de una exposición comercial en Las Vegas que había durado una semana. CONEXPO era la exposición sobre construcción más grande del mundo y se hacía sólo cada cinco años; atraía miles de exhibidores y participantes. Todo aquel que tuviera algo que ver con la construcción se encontraba ahí. Y por primera vez, HH Industries había estado justo en el centro de todo.

"Nuestro local de exhibición estaba increíble", continuó Hal. "No tengo idea de lo que John le prometió a la chica encargada de organizar aquello, pero ¡valió la pena!", le hizo un guiño a Laurel. "Definitivamente obtuvimos una ubica-

ción excelente. Saliendo del ascensor, tenías que pasar justo delante de nuestro local para llegar a la parte principal de la sala de exhibiciones. Y lo que diseñamos para exhibir de veras que llamaba la atención. No sólo repartimos los 400 catálogos que llevábamos, sino que tenemos direcciones de al menos otro tanto que desea estar incluido en nuestra lista de correos."

Laurel sonrió, el entusiasmo de Hal era contagioso. "Habías dicho que la exhibición podría muy bien cambiar el futuro de nuestra empresa, y parece que tenías razón. Stan dijo que incluso comieron con clientes todas las tardes. Apuesto a que te alegras de estar de regreso, a pesar de que sea solamente para tomar aire."

"Lo mejor es que ya terminó", concordó Hal. "Sólo espero que nuestra inversión tenga frutos; en total, gastamos casi $15,000 para participar."

Laurel se quedó pensativa unos momentos. "Las cifras de las ventas podrían tardar un buen rato en reaccionar con la fuerza suficiente para que lo notemos. La gente no cambia de proveedores tan pronto. Sin embargo, tengo una idea de cómo podemos medir una respuesta más inmediata. ¿No tiene Stan un registro del número de clientes distintos que nos llaman todos los días?"

Hal asintió sonriendo. "La CUC, cuenta única del cliente. Ya conoces a Stan, ¡mantiene registros de todo!"

"Eso pensé. Déjame echar una mirada a los datos de las CUC de las últimas semanas. Luego recolectamos algunos datos más, empezando el lunes. Podemos comparar las cifras antes y después de CONEXPO, y determinar si vemos alguna diferencia en la cantidad de personas con las que tenemos comunicación. Es una técnica sencilla, pero puede darnos una idea en el corto plazo. Más adelante podemos desarrollar un análisis más detallado y profundo para rastrear las tendencias en las ventas, proyectar un punto de equilibrio para nuestra inversión, etcétera."

"Me parece bien", comentó Hal, "déjame saber qué obtienes".

Laurel recolectó los siguientes datos (que también puede consultar en los archivos CH09A.xxx del CD que acompaña al libro):

CUC PreCONEXPO	CUC PosCONEXPO
39 22 28 11 33 47 25	41 47 27 27 36 24 28
24 34 28 15 40 27 33	36 46 34 36 44 13 23
27 23 27 18 46 33 17	46 48 48 44 45 26 36
38 38 19 13 30 20 38	21 30 35 36 23 36 34
37 33 17 28 36 23 27	36 48 38 44 35 36 46
34 24 38 35 10 37 23	36 36 24 26 24 23 23
29 32 38 21 43 27 35	51 37 37 17 46 34 28
28 18 39 23 20 30 28	41 30 25 46 33 26 44
17 33 47 28 16 23 37	36 38 28 34 45 46 36

1. Determine si las poblaciones de llamadas PreCONEXPO y PosCONEXPO tienen el mismo número medio de clientes únicos por día. Pruebe con un nivel de significancia de 0.05.

Como era de esperarse, Hal estaba animado con los hallazgos de Laurel. "Parece que vamos a tener algunos clientes nuevos", comentó en la siguiente reunión de directivos. John Raymond miró a Laurel, que estaba sentada al otro extremo de la mesa. "¿Laurel, hay alguna manera de poder saber si nos perjudicó no haber tenido suficientes catálogos en la exposición? Quiero decir, ¿podríamos saber si los clientes que recibieron un catálogo por correo nos compran menos que los que obtuvieron uno en nuestro puesto?.

Stan levantó la mirada. "Eso sería estupendo", intervino, "y como durante la exposición llamamos acá para pedir más catálogos, los clientes que pidieron uno y no pudimos dárselo allá, debieron tenerlo en cuanto llegaron a su lugar de origen; tal vez podamos, incluso, utilizar los datos del mismo periodo para este estudio", concluyó señalando los papeles que tenía enfrente.

Laurel meditó unos momentos. "Eso será un buen reto", empezó. "Los tamaños de población son bastante diferentes, de modo que tendríamos que usar un enfoque de proporciones, en lugar de las cifras de ventas reales. Déjame ver qué puedo obtener. Stan, necesito que me ayudes a seleccionar a los clientes relevantes en los datos de las últimas semanas. Pasaré a tu oficina por la tarde."

Laurel empezó a diseñar el estudio tan pronto como salió de la junta de directivos. Stan había marcado los informes de ventas, indicándole por separado los clientes pertenecientes a cada grupo. Luego ella calculó el porcentaje de clientes que hicieron pedidos cada día durante el periodo de estudio. (Estos datos están en los archivos CH09B.xxx del CD que acompaña al libro.)

Catálogo recogido

0.019	0.081	0.003	0.052	0.092	0.032	0.021
0.079	0.019	0.014	0.061	0.023	0.007	0.023
0.026	0.017	0.074	0.018	0.072	0.023	0.038
0.022	0.061	0.031	0.025	0.020	0.027	0.029
0.080	0.037	0.021	0.027	0.024	0.057	0.067
0.036	0.091	0.025	0.056	0.046	0.086	0.042
0.028	0.010	0.005	0.056	0.009	0.024	0.061
0.019	0.010	0.021	0.057	0.017	0.017	0.057
0.015	0.054	0.022	0.019	0.031	0.053	0.041

Catálogo enviado por correo

0.016	0.095	0.011	0.002	0.026	0.011	0.055
0.030	0.030	0.051	0.017	0.048	0.050	0.021
0.003	0.046	0.012	0.014	0.072	0.013	0.056
0.052	0.009	0.085	0.024	0.015	0.022	0.045
0.028	0.012	0.013	0.015	0.013	0.027	0.049
0.014	0.061	0.007	0.023	0.023	0.007	0.035
0.019	0.030	0.029	0.042	0.022	0.065	0.033
0.012	0.052	0.019	0.030	0.041	0.014	0.032
0.074	0.023	0.027	0.057	0.056	0.017	0.053

2. Determine si la proporción promedio diaria de clientes que hacen pedidos es menor para quienes recibieron el

catálogo por correo (comparado con los que lo obtuvieron en la exposición). ¿Para qué valores de α se acepta la hipótesis nula?

Stan estaba complacido con los resultados obtenidos hasta ese entonces, pero Laurel sabía que no estaba del todo satisfecho. "Ya sé qué necesitas", le dijo Laurel una tarde después de una larga discusión. "¡Necesitas dinero! Nunca estás completamente contento si no hay dólares de por medio."

Stan sonrió tímidamente. "Me conoces demasiado bien. Tengo curiosidad de saber, en el fondo, cómo nos afectó nuestra participación en CONEXPO. A mi modo de ver, las cifras de ventas están bien, pero tú y tu estadística deberán convencerme de que no se trata únicamente de mis buenos deseos."

Stan provenía definitivamente de la vieja escuela en que la estadística era un asunto de académicos y no de hombres de negocios. Había tardado mucho en aceptar que el trabajo estadístico era válido y valioso para HH Industries. "Tomemos unas 40 unidades del inventario y estudiemos su desempeño en ventas durante el mismo periodo en que realizamos los estudios de las cuentas únicas y los catálogos", dijo ella. "Eso nos dará una buena idea de cómo están las ventas en general." Con la ayuda de Stan, Laurel recolectó los siguientes datos para las mismas 40 unidades, antes y después de la CONEXPO. (Estos datos están en los archivos CH09B.xxx del CD que acompaña al libro.)

Datos sobre ventas antes de CONEXPO

1,163	711	394	538	915
850	598	1,619	1,624	486
766	670	469	2,125	860
305	609	1,510	512	782
1,800	818	531	521	1,561
318	1,215	910	1,441	719
360	834	1,195	530	321
266	225	925	1,762	652

Datos sobre ventas después de CONEXPO

1,220	795	316	713	814
932	663	1,801	1,848	538
810	891	471	2,118	888
315	812	1,620	659	950
2,015	930	609	581	1,691
325	1,422	934	1,493	842
500	850	1,300	695	419
358	252	1,287	1,873	775

3. A un nivel de significancia del 15%, determine si las ventas medias se incrementaron después de CONEXPO. ¿Se puede atribuir por completo el efecto sobre las ventas a la CONEXPO? Explique su respuesta.

Del libro de texto al mundo real

Los ensayos clínicos utilizan la estadística para determinar los mejores procedimientos médicos

Los médicos a menudo bromean sobre el hecho de que cuando una persona está enferma, uno de los lugares más peligrosos en que puede estar es un hospital. La broma se refiere a que, en ocasiones, los pacientes de un hospital se contagian de otros pacientes. Una investigación ha mostrado que la mayor parte de los contagios ocurren a través de las manos de los trabajadores de la salud, al pasar de un paciente a otro. Así, un grupo de médicos y enfermeras del Colegio de Medicina de la Universidad de Iowa decidieron investigar la práctica del lavado de manos en las tres unidades de cuidado intensivo (UCI), en un periodo de ocho meses.

El doctor Bradley Doebbling y sus colegas establecieron un estudio cruzado para comparar dos tipos de jabones. Cada mes cambiaban el jabón disponible. Ambos jabones tienen algún agente antibacteriano: uno contiene el antibiótico gluconato de clorhexidina y el otro contiene una solución con 60% de alcohol.

El hospital tiene ya establecidos procedimientos para la detección de infecciones en los pacientes de las UCI y su rápido tratamiento. Los investigadores también han establecido periodos de observación aleatoria cada media hora para registrar el número de veces que los médicos y las enfermeras de las UCI se lavan las manos entre revisiones. Después compararon el número de infecciones reportadas bajo los dos sistemas de lavado de manos, utilizando el número de *días paciente* bajo los dos regímenes (un paciente que esté en el hospital durante un día cuenta como un *día paciente*).

Agente limpiador de manos

	Clorhexidina	Alcohol
Número de días paciente	4,001	3,984
Número de infecciones	152	202

Demostración de que hay diferencia entre limpiadores Los resultados obtenidos fueron que durante los meses en que se utilizó la clorhexidina, la frecuencia de infecciones en el hospital fue sólo las tres cuartas partes de cuando se usó la solución de alcohol; un análisis estadístico de la tasa de infecciones de los dos grupos indicó que esta diferencia es significativa.

La aparente efectividad superior de la clorhexidina depende, en parte, de la disposición de los profesionales para cumplir las indicaciones de lavado de manos. En general, los investigadores encontraron que los miembros del personal clínico se lavaban las manos en sólo alrededor del 40% de las ocasiones en que hubiera ayudado al control de infecciones, pero estaban más dispuestos a lavarse cuando tenían clorhexidina. En una de las tres unidades de cuidado intensivo, la diferencia en el cumplimiento del lavado de manos llegó al 48% con la clorhexidina contra el 30% con el uso de la solución de alcohol ($p = 0.002$ con una prueba t).

El beneficio es sustancial Los estudios han estimado el costo anual total del tratamiento de infecciones en hospitales, en Estados Unidos, entre 5 y 10 mil millones de dólares. Exhortar al personal médico de un hospital a lavarse bien las manos entre revisiones parece disminuir mucho las tasas de infecciones, y es más probable que el personal se lave las manos cuando tienen al alcance un limpiador que contenga un fuerte agente antibacteriano.

Fuente: Doebbling, B. N., G. L. Stanley, C. T. Sheetz, M. A. Pfaller, A. K. Houston, L. Annis, N. Li, y R.P. Wenzel. "Comparative Efficacy of Alternative Hand-Washing Agents in Reducing Nosocomial Infections in Intensive Care Units." *The New England Journal of Medicine*, 1992, vol. 327(2), págs. 88-93.

Repaso del capítulo

● Términos introducidos en el capítulo 9

Estimación conjunta de σ^2 Promedio ponderado de σ_1^2 y σ_2^2, usado para estimar la varianza común, σ^2, al usar muestras pequeñas para probar la diferencia entre dos medias de población.

Muestras apareadas Otro nombre para las muestras dependientes.

Muestras dependientes Muestras extraídas de dos poblaciones de tal manera que los elementos de una muestra se aparean con los elementos de la otra muestra, con el fin de permitir un análisis más preciso al controlar los factores externos.

Proporción combinada de éxitos Al comparar dos proporciones de población, el número total de éxitos en ambas muestras dividido entre el tamaño total de ambas muestras se usa para estimar la proporción de éxitos comunes a ambas poblaciones.

Prueba de diferencias apareadas o por pares Una prueba de hipótesis de la diferencia entre dos medias de población basada en las medias de dos muestras dependientes.

Pruebas de dos muestras Pruebas de hipótesis basadas en muestras tomadas de dos poblaciones con el fin de comparar sus medias o proporciones.

Valor P Otro nombre para un valor prob y es el mayor nivel de significancia en el que aceptaríamos la hipótesis nula. Nos permite probar hipótesis sin especificar primero un valor para α.

● Ecuaciones introducidas en el capítulo 9

■ **9-1**
$$\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Esta fórmula nos permite derivar la desviación estándar de la distribución de la diferencia entre dos medias muestrales, esto es, *el error estándar de la diferencia entre dos medias*. Para hacer esto, tomamos la raíz cuadrada de la suma de la varianza de la población 1 dividida entre el tamaño de su muestra y la varianza de la población 2 dividida entre el tamaño de su muestra.

■ **9-2**
$$\hat{\sigma}_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}$$

Si se desconocen las desviaciones estándar de las dos poblaciones, podemos usar esta fórmula para derivar el error estándar *estimado* de la diferencia entre dos medias. Podemos usar esta ecuación después de haber utilizado las dos desviaciones estándar de la muestra y la ecuación 7-1 para determinar las desviaciones estándar estimadas de la población 1 y población 2 ($\hat{\sigma} = s$).

■ **9-3**
$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Con esta fórmula podemos obtener una "estimación conjunta" de σ^2. Utiliza un promedio ponderado de s_1^2 y s_2^2, donde los pesos son los grados de libertad en cada muestra. El uso de esta fórmula supone que

$\sigma_1^2 = \sigma_2^2$ (las varianzas de población desconocidas son iguales). Utilizamos esta fórmula cuando hacemos una prueba de las diferencias entre medias en situaciones en que se tienen tamaños de muestra pequeños (menores que 30).

■ **9-4**
$$\hat{\sigma}_{\bar{x}_1 - \bar{x}_2} = s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Dada la "estimación conjunta" de α^2 obtenida con la ecuación 9-3, sustituimos este valor en la ecuación 9-2 y simplificamos la expresión. Esto nos da una fórmula para estimar el error estándar de la diferencia entre medias muestrales en el caso en que tengamos muestras pequeñas (menores que 30), pero varianzas de población iguales.

■ **9-5**
$$\sigma_{\bar{p}_1 - \bar{p}_2} = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

Ésta es la fórmula que se utiliza para derivar el error estándar de la diferencia entre dos *proporciones*. Los símbolos p_1 y p_2 representan las proporciones de éxitos en las poblaciones 1 y 2 respectivamente, y q_1 y q_2 son las proporciones de fracasos en las poblaciones 1 y 2, respectivamente.

■ **9-6**
$$\hat{\sigma}_{\bar{p}_1 - \bar{p}_2} = \sqrt{\frac{\bar{p}_1 \bar{q}_1}{n_1} + \frac{\bar{p}_2 \bar{q}_2}{n_2}}$$

Si no se conocen los parámetros de la población p y q, podemos utilizar los estadísticos de la muestra $\bar{p}$ y $\bar{q}$, y esta fórmula para *estimar* el error estándar de la diferencia entre dos proporciones.

■ **9-7**
$$\hat{p} = \frac{n_1 \bar{p}_1 + n_2 \bar{p}_2}{n_1 + n_2}$$

Debido a que la hipótesis nula supone que *no hay diferencia* entre las dos proporciones de población, sería más apropiado modificar la ecuación 9-6 y utilizar las proporciones combinadas de ambas muestras para estimar la proporción global de éxitos en las poblaciones combinadas. La ecuación 9-7 combina las proporciones de ambas muestras. Note que el valor de q es igual a $1 - \hat{p}$.

■ **9-8**
$$\hat{\sigma}_{\bar{p}_1 - \bar{p}_2} = \sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}}$$

Ahora ya podemos sustituir los resultados de la ecuación 9-7, tanto $\hat{p}$ como $\hat{q}$, en la ecuación 9-6 y obtener una versión más correcta. Esta nueva ecuación, 9-8, nos da el error estándar *estimado* de la diferencia entre las dos proporciones, utilizando estimaciones combinadas de ambas muestras.

Ejercicios de repaso

■ **9-38** La fábrica Clic Pens ha probado dos tipos de anuncios en puntos de venta para su nueva pluma de tinta borrable. El anuncio de mostrador fue colocado en una muestra aleatoria de 40 tiendas en el mercado de prueba, y el anuncio de piso se colocó en otras 40 tiendas del área. El número medio de plumas vendidas por tienda en un mes con los anuncios de mostrador fue 42, con una desviación estándar de 8. En cuanto a los anuncios de piso, el número medio de plumas vendidas por tienda durante el mismo mes fue 45, con desviación estándar de 7. Al nivel $\alpha = 0.02$, ¿hubo una diferencia significativa entre las ventas con los dos tipos de anuncio?

■ **9-39** En 1992, una investigación en 50 hospitales municipales reveló una tasa de ocupación media del 73.6%, con una desviación estándar de la muestra del 18.2%. Otra investigación efectuada en 75 hospitales municipales en 1995, encontró una tasa de ocupación promedio del 68.9%, con una desviación estándar de la muestra del 19.7%. A un nivel $\alpha = 0.10$, ¿podemos concluir que la tasa de ocupación promedio cambió significativamente durante los 3 años que hay entre las dos investigaciones?

■ **9-40** La compañía General Cereals acaba de concluir una nueva campaña publicitaria para su producto Fruit Crunch, el cereal natural para el desayuno que contiene nueces, granos y fruta seca. Para probar la efec-

tividad de la campaña, el gerente de marca, Alan Neebe, encuestó a 11 clientes antes de la campaña y a otros 11 después. A continuación se da el consumo semanal (en onzas) de Fruit Crunch por parte de los consumidores:

Antes	14	5	18	18	30	10	8	26	13	29	24
Después	23	14	13	29	33	11	12	25	21	26	34

a) Al nivel $\alpha = 0.05$, ¿puede Alan concluir que la campaña ha tenido éxito en aumentar la demanda del producto?
b) Dada la investigación de Alan antes de la campaña, ¿puede usted sugerirle un mejor procedimiento de muestreo para después de la campaña?

■ **9-41** La compañía Ben & Jerry's Homemade, es una empresa que se dedica a la venta de helados no convencionales, con sabores extravagantes como el de galleta con chispas de chocolate. En un artículo en *The Wall Street Journal* se consigna que parte del éxito de la compañía se debe a que atrae a los adultos jóvenes (quienes presumiblemente serán fieles a la compañía durante su periodo pico de consumo de helados). Suponga que un investigador de mercado lleva a cabo una encuesta a la salida de un supermercado y hace una sola pregunta a 200 compradores consecutivos del helado Ben & Jerry's y a 200 compradores de la marca competidora de Haagen-Daz de Grand Metropolitan: "¿Tiene usted menos de 25 años?" Si el 7% de los aficionados al helado de Ben & Jerry's dice "sí" y solamente el 3% de los aficionados al otro helado dice "sí", ¿esto confirma la conclusión del artículo, para 10% de nivel de significancia?

Fuente: Suein L. Hwang, "While Many Campetitors See Sales Melt, Ben & Jerry's Scoops Out Solid Market Growth" (Mientras muchos competidores ven derretirse sus ventas, Ben & Jerry's construye un sólido crecimiento), en *The Wall Street Journal* (25 de mayo de 1993): B1.

■ **9-42** Students Against Drunk Driving ha establecido que el uso del cinturón de seguridad es un paso positivo para reducir accidentes y lesiones. Antes de una campaña importante en una secundaria, el 44% de 150 conductores que entraba al estacionamiento de la escuela llevaban puesto el cinturón. Después del programa de sensibilización, la proporción de uso se elevó al 52% en una muestra de 200 vehículos. A un nivel de significancia de 0.04, ¿pueden los estudiantes concluir que su campaña fue efectiva?

■ **9-43** La compañía distribuidora Allen tiene la hipótesis de que una llamada telefónica es más efectiva que una carta para acelerar el pago de cuentas atrasadas. Se contactaron dos grupos de cuentahabientes atrasados, cada uno con uno de los dos métodos, y se registró el lapso entre el envío de la carta o la llamada y el momento en que se recibió el pago:

Método utilizado	**Días hasta el pago**						
Carta	10	8	9	11	11	14	10
Llamada telefónica	7	4	5	4	8	6	9

a) Al nivel $\alpha = 0.025$, ¿debería Allen concluir que las cuentas atrasadas se pagan más pronto con una llamada que con una carta?
b) ¿Puede Allen concluir que las cuentas atrasadas responden con mayor rapidez a las llamadas telefónicas?

■ **9-44** Recientemente, un analgésico perdió parte de sus ventas ante un nuevo competidor. El competidor anunciaba que su marca entra al torrente sanguíneo con más rapidez que el *otro* analgésico y, como resultado, alivia el dolor más pronto. El laboratorio que fabrica el primer analgésico desea probar que no existe diferencia significativa entre los dos productos y, por consiguiente, que la afirmación de su competidor es falsa. Como prueba preliminar, se suministró el primer analgésico a 9 individuos una vez al día durante 3 semanas. En las siguientes 3 semanas, se proporcionó a los mismos sujetos el producto de la competencia. Para cada medicamento, se registró el tiempo promedio en minutos que les tomó aliviar el dolor:

Individuo	1	2	3	4	5	6	7	8	9
Analgésico	16.5	25.5	23.0	14.5	28.0	10.0	21.5	18.5	15.5
Competidor	12.0	20.5	25.0	16.5	24.0	11.5	17.0	15.0	13.0

A un nivel $\alpha = 0.10$, ¿existe alguna diferencia significativa en el tiempo que tomó a los medicamentos aliviar el dolor?

■ **9-45** Considere los datos de la tabla que sigue. En los primeros tres meses de 1993, la venta de automóviles de lujo en Estados Unidos disminuyó ligeramente, pero la proporción de automóviles de lujo que fueron importados se incrementó. Las cifras se muestran en la siguiente tabla. ¿Es significativo el cambio en el porcentaje de mercado de los automóviles importados con respecto a los domésticos, al nivel del 5%?

	Ventas de autos de lujo 1992	Primer trimestre 1993
Porcentaje de autos nacionales	47.5	46.2
Porcentaje de autos importados	52.5	53.8
Total de unidades vendidas	373,842	372,442

Fuente: Business Week *(31 de mayo de 1993), pág. 40.*

■ **9-46** Un químico que se encuentra desarrollando repelentes de insectos desea saber si una nueva fórmula da mayor protección contra los piquetes que la que proporciona el producto líder del mercado. En un experimento, se roció un brazo de 14 voluntarios con el producto viejo y el otro brazo con la nueva fórmula. Luego cada individuo metió los brazos en una cámara llena de mosquitos y otros insectos que pican. El número de piquetes recibidos en cada brazo se registró en la tabla siguiente. Al nivel $\alpha = 0.01$, ¿deberá el químico concluir que la nueva fórmula, de hecho, es más efectiva que el producto líder en el mercado?

Individuo	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Fórmula vieja	5	2	5	4	3	6	2	4	2	6	5	7	1	3
Fórmula nueva	3	1	5	1	1	4	4	2	5	2	3	3	1	2

■ **9-47** Long Distance Carrier (LDC) intenta observar el efecto de ofrecer "1 mes gratis" con una cuota fija de $10.95, contra la oferta de una cuota mensual más baja ($8.75) sin el mes gratis. Para probar qué puede ser más atractivo para los consumidores, Long Distance realiza una prueba de mercado breve: 12 representantes hacen llamadas usando un enfoque y 10 usan el otro. El siguiente número de clientes estuvo de acuerdo en cambiar su servicio de larga distancia actual a LDC:

Oferta	**Número de cambios**											
1 mes gratis	118	115	122	99	106	125	102	100	92	103	113	129
Cuota mensual baja	115	126	113	110	135	102	124	137	108	128		

Pruebe para un nivel de significancia del 10% si existen diferencias significativas entre las dos ofertas.

■ **9-48** ¿El nivel percibido de responsabilidad por una acción está relacionado con la severidad de sus consecuencias? Esta pregunta fue la base de un estudio sobre responsabilidad en el que los sujetos leían la descripción de un accidente ocurrido en una carretera federal. Las consecuencias, en términos de costos y daños físicos, se describían como menores y graves. Se utilizó un cuestionario para calificar el grado de responsabilidad que los individuos creían que debían asignar al personaje principal. A continuación se dan los resultados para los grupos de consecuencias menores y consecuencias graves. Las calificaciones altas corresponden a una responsabilidad mayor atribuible al personaje principal. Si se utiliza un nivel de significancia del 0.025, ¿se concluye del estudio que las consecuencias graves conducen a una mayor atribución de responsabilidad?

Consecuencias	**Grado de responsabilidad**							
Menores	4	5	3	3	4	1	2	6
Graves	4	5	4	6	7	8	6	5

■ **9-49** En octubre de 1992, una investigación entre 120 macroeconomistas indicó que 87 de ellos creían que la recesión ya había terminado. Una investigación de 150 agentes de adquisiciones encontró que 89 de éstos creían que la recesión ya había terminado. Al nivel $\alpha = 0.10$, ¿se pude concluir que los agentes de adquisiciones eran más pesimistas acerca de la economía de Estados Unidos que los macroeconomistas?

■ **9-50** Los largometrajes de animación proporcionan grandes ganancias a The Walt Disney Company. Las películas de éxito, como *Aladdin*, pueden producir ingresos mayores a los 300 millones de dólares después de sus costos de producción. Pero cada año, además de estos éxitos espectaculares, algunas películas de animación más modestas llegan a las pantallas. Un analista obtuvo un listado de los ingresos totales (la suma de boletos vendidos) de la primera presentación de todas las películas de animación exhibidas entre 1986 y 1992, los datos reunidos son los siguientes (en millones de dólares):

Película	**Estudio**	**Ingreso total (millones de dólares)**
Un Cuento Americano	Universal	44.9
Los Perros Van al Cielo	MGM/UA	26.2
Ferngully: El Bosque Perdido	Twentieth Century Fox	20.9

Película	Estudio	Ingreso total (millones de dólares)
Fievel va al Oeste	Universal	20.2
Mundo Cool	Paramount	13.7
Rock-a-Doddle	Twentieth Century Fox	11.6
Bebe's Kids	Paramount	7.5
Ositos Cariñositos II	Columbia	5.4
Pinocho y el Emperador de la Noche	New World	2.7
Ositos Cariñositos	Cineplex	2.2
Babar: la Película	New Line	1.4
Oliver y Compañía	Disney	52.6
Rescuers Down Under	Disney	27.8
El Gran Ratón Detective	Disney	24.2
Los Rescatadores	Disney	21.1
Patoaventuras	Disney	18.1

(Nota: Se excluyen los éxitos de Disney que hayan tenido ganancias mayores que 80 millones de dólares y los reestrenos de los clásicos animados como el de Blancanieves.)

Fuente: David J. Londoner y William L. Kessler, The Walt Disney Company: Momentum in the Basics, Wertheim Schroeder & Co. Informe de investigación, Nueva York, 12 de febrero de 1993.

Para $\alpha = 0.05$, ¿estos datos apoyan la conclusión de que "aun excluyendo los grandes éxitos, las películas animadas de Disney ganan más que la competencia?".

■ **9-51** El programa de estudios de licenciatura en administración de Piedmont Business School ofrece un Taller de Habilidades Analíticas (THA), durante el verano, para ayudar a los estudiantes a refrescar sus conocimientos de contabilidad, economía y matemáticas. El director del programa, Andy Bunch, se pregunta si el THA ha proporcionado ventajas a quienes lo cursan. Ha tomado muestras aleatorias de los promedios de los asistentes durante los últimos 5 años y de los que iniciaron la maestría en administración sin el THA durante el mismo periodo. A un nivel $\alpha = 0.02$, ¿han obtenido los estudiantes que cursaron el THA calificaciones significativamente más altas? ¿Debería Andy anunciar que el taller ayuda a los estudiantes a lograr mejores resultados en su programa de maestría?

	$\bar{x}$	s	n
Alumnos con THA	3.37	1.13	26
Alumnos sin THA	3.15	1.89	35

■ **9-52** De 2,000 empresas muestreadas aleatoriamente, 58 tuvieron auditorias respecto a los reembolsos correspondientes al pago de impuestos federales de 1995, en Estados Unidos. En otra muestra de 2,500 empresas, 61 tuvieron en 1994 auditorias respecto a sus reembolsos, ¿la fracción de reembolsos corporativos que tuvieron auditoria en 1995 fue significativamente distinta de la fracción de reembolsos de 1994? Pruebe las hipótesis apropiadas a un nivel $\alpha = 0.0$l.

■ **9-53** Ellen Singer aseguró a uno de sus compañeros de trabajo que las casas de la parte sur del condado de Durham se venden en aproximadamente $15,000 menos que las casas del mismo tipo que se venden en Chapel Hill. Para probar esta afirmación, su compañero escogió al azar 10 ventas recientes de casas en Chapel Hill y las comparó con 10 ventas de recientes en el condado de Durham, en términos de estilo, tamaño, edad, número de cuartos y tamaño del terreno que ocupan. A un nivel $\alpha = 0.05$, ¿los datos siguientes (precios de ventas en miles de dólares) apoyan la afirmación de Ellen?

Chapel Hill	97.3	108.4	135.7	142.3	151.8	158.5	177.4	183.9	195.2	207.6
Condado de Durham	81.5	92.0	115.8	137.8	150.9	149.2	168.2	173.9	175.9	194.4

■ **9-54** El Promedio de Transporte Dow-Jones está basado en los precios de cierre de las acciones comunes de 20 compañías aéreas, ferrocarriles y de fletes. El 24 de mayo de 1993, el precio promedio de estas 20 acciones disminuyó de $47.156 a $47.150. Tomando estas acciones como una muestra aleatoria de todas las acciones relacionadas con el transporte, ¿es significativa la disminución observada? Explique su respuesta.

Acciones	Cierre 5/21/93	Cierre 5/24/93
Airborne Freight	23.500	23.375
Alaska Airlines	16.750	16.625

Acciones	Cierre 5/21/93	Cierre 5/24/93
American President	50.750	52.000
AMR	71.625	71.875
Burlington Northern	53.875	54.000
Carolina Freightways	13.125	13.000
Conrail	51.375	51.875
Consolidated Freightways	15.375	15.625
CSX	71.125	70.750
Delta Airlines	60.250	60.750
Federal Express	49.750	49.375
Norfolk and Southern	61.875	61.250
Roadway	54.750	55.000
Ryder System	27.250	27.125
Santa Fe Pacific	16.750	17.000
Southwest Airline	40.375	40.500
UAL	138.750	138.750
Union Pacific	63.250	63.000
USAir Group	22.500	22.625
XTRA	40.125	38.500
Average Price	47.156	47.150

Fuente: The Wall Street Journal *(25 de mayo de 1993): C3-C6.*

■ **9-55** La ejecutiva de la televisión, Terri Black, acaba de recibir una propuesta y una cinta piloto para una nueva serie. *El nido ya no está vacío* es una comedia sobre un matrimonio de mediana edad cuyos hijos, dos graduados universitarios, son incapaces de encontrar trabajo y regresan a vivir con sus padres. Terri se pregunta si el programa llamará la atención tanto de los televidentes con veintitantos años como de los mayores. Calculando que las personas que trabajan en su oficina son una muestra razonablemente representativa del total de la población, les pide que hagan una evaluación de la cinta en una escala de 0 a 100 y obtiene la siguiente respuesta:

Edad	Respuesta						
20-29	86	74	73	65	82	78	79
≥ 30	63	72	68	75	73	80	

a) A un nivel de significancia de 0.05, ¿deberá Terri concluir que el programa es igualmente atractivo para los dos grupos?

b) Independientemente de su respuesta al inciso anterior, ¿piensa que Terri debe utilizar los resultados de la investigación que hizo en su oficina para decidir cómo diseñar una campaña publicitaria para *El nido ya no está vacío*? Explique su respuesta.

■ **9-56** Un fabricante de alimentos para mascotas se pregunta si las personas que tienen gatos y las que tienen perros reaccionan de manera diferente respecto a los alimentos que fabrica. Se llevó a cabo una investigación entre los consumidores que arrojó la siguiente información:

Tipo de mascota	Dueños entrevistados	Cantidad que usa el alimento
Gato	280	152
Perro	190	81

¿Es razonable concluir a un nivel $\alpha = 0.02$, que es más probable que los dueños de gatos alimenten a su mascota con el alimento en cuestión?

■ **9-57** Ofrecieron a Robin Wendell transferirlo de Pittsburgh a Boston, pero él quiere un sueldo más alto "porque la vida es mucho más cara allá". De su recibo de supermercado quitó los artículos costosos y obtuvo una lista de 36 artículos de menos de $2 con media de $0.98 y desviación estándar de $0.43 en Pittsburgh. El gerente de personal, en Boston va al supermercado y con el mismo límite de $2 compra 42 artículos, con media de $1.07 y desviación estándar de $0.38. ¿Tiene razón Robin en cuanto a que el costo de los

abarrotes es mayor en Boston que en Pittsburgh para un nivel de confianza de $\alpha = 0.01$? ¿Qué puede hacer para mejorar el análisis del costo de vida en las dos ciudades?

■ **9-58** Un grupo de médicos ingleses repitieron un estudio administrativo en Estados Unidos acerca de directores médicos que preguntaba qué habilidades consideraban útiles y esenciales para sus trabajos. De 100 encuestas enviadas por correo en Estados Unidos, respondieron el 50% y de 59 encuestas enviadas en el Reino Unido, regresaron el 45%.

a) ¿Cuántos cuestionarios de la encuesta del Reino Unido regresaron?
b) ¿Difieren las proporciones de respuesta entre las dos muestras?

Fuente: Trevor Wood, Gabriel Scally y Declan O'Neill, "Management Knowledge and Skills Required by UK and US Medical Directors", *Physician Executive* (agosto de 1995): 26-29.

■ **9-59** Una compañía está considerando ofrecer a un nuevo gerente con maestría en administración una ayuda de renta como compensación por la diferencia en el costo de la vivienda. Un ejecutivo hace referencia a una tabla en *The Wall Street Journal* que contiene un estudio de precios típicos de casas para jóvenes ejecutivos que inician su carrera, en todo el país. El ejecutivo observa, "Los precios de alojamiento están bajando en California; debemos sólo compensar a las personas que viven en la costosa costa este y en el sur".

Este y sur	Medio oeste y oeste
159,125	167,155
206,125	95,000
188,071	197,519
196,750	215,368
143,792	145,767
324,171	154,914
248,157	153,581
199,419	156,690
110,650	178,169
169,586	210,914
174,800	

a) Si la tabla representa una muestra aleatoria de los precios de alojamiento en cada área, ¿el precio medio es significativamente diferente entre los dos grupos de vivienda?
b) Al avanzar la discusión, un colega indica que la cifra de $324,171 en el primer grupo corresponde a Washington, D.C., y es evidente que es una observación extrema. La persona de recursos humanos aclara que la cifra de $95,000 del segundo grupo es de Eden Prairie, Minnesota, y dice, "ni siquiera tenemos oficinas en Minnesota". ¿Mejora este análisis si se eliminan estos valores extremos? ¿Se llega a las mismas conclusiones?

capítulo 11

JI-CUADRADA Y ANÁLISIS DE VARIANZA

Objetivos

- Reconocer situaciones que requieren la comparación de más de dos medias o proporciones
- Introducir las distribuciones ji-cuadrada y *F*, y aprender a usarlas en inferencia estadística
- Utilizar la distribución ji-cuadrada para ver si dos clasificaciones de los mismos datos son independientes entre sí
- Utilizar una prueba ji-cuadrada para probar si una colección particular de datos está bien descrita por una distribución especificada
- Utilizar la distribución ji-cuadrada para intervalos de confianza y prueba de hipótesis respecto a una sola varianza de población
- Comparar más de dos medias de población empleando el análisis de varianza
- Utilizar la distribución *F* en pruebas de hipótesis de dos varianzas de población

Contenido del capítulo

11.1 Introducción 448
11.2 Ji-cuadrada como prueba de independencia 449
11.3 Ji-cuadrada como prueba de bondad de ajuste: prueba de lo apropiado de una distribución 462
11.4 Análisis de varianza 468
11.5 Inferencias acerca de una varianza de población 484
11.6 Inferencias acerca de las varianzas de dos poblaciones 489

- Estadística en el trabajo 496
- Ejercicio de base de datos computacional 496
- Del libro de texto al mundo real 498
- Términos introducidos en el capítulo 11 498
- Ecuaciones introducidas en el capítulo 11 499
- Ejercicios de repaso 501

El director de capacitación de una compañía está evaluando tres métodos de capacitación para empleados nuevos. El primero consiste en asignar un empleado nuevo a un trabajador experimentado para que éste le ayude en la fábrica; el segundo, en ubicar a todos los empleados nuevos en un salón de capacitación separado de la fábrica; el tercer método utiliza películas y materiales de aprendizaje programados. El director de capacitación escoge al azar 16 empleados nuevos asignados al azar a los tres métodos, registra su producción diaria después de terminar los programas de capacitación:

Método 1	**15**	**18**	**19**	**22**	**11**	
Método 2	**22**	**27**	**18**	**21**	**17**	
Método 3	**18**	**24**	**19**	**16**	**22**	**15**

El director se pregunta si existen diferencias en la efectividad de los tres métodos. Con las técnicas descritas en este capítulo, podemos ayudarle a responder esa pregunta. ■

11.1 Introducción

En los capítulos 8 y 9 aprendimos a probar hipótesis utilizando datos provenientes de una o dos muestras. Usamos pruebas de una muestra para determinar si una media o una proporción era significativamente diferente de un valor hipotético. En las pruebas de dos muestras examinamos la diferencia entre dos medias o entre dos proporciones e intentamos averiguar si esta diferencia era significativa.

Usos de la prueba ji-cuadrada

Suponga que tenemos proporciones de cinco poblaciones en lugar de solamente dos. En este caso, los métodos para comparar proporciones, descritos en el capítulo 9, *no* son pertinentes; debemos utilizar la *prueba ji-cuadrada*, que es el tema de la primera parte del presente capítulo. Las pruebas ji-cuadrada nos permiten probar si *más* de dos proporciones de población pueden ser consideradas iguales.

En realidad, las pruebas ji-cuadrada nos permiten hacer mucho más que probar la igualdad de varias proporciones. Si clasificamos una población en diferentes categorías respecto a dos atributos (por ejemplo, edad y desempeño en el trabajo), entonces podemos utilizar una prueba ji-cuadrada para determinar si los dos atributos son independientes entre sí.

Función del análisis de varianza

Los administradores también se topan con situaciones en las que resulta útil probar la igualdad de más de dos medias de población. De nuevo: no podemos aplicar los métodos introducidos en el capítulo 9 debido a que están limitados a la prueba de la igualdad de solamente dos medias. El análisis de varianza, que se estudia en la cuarta sección de este capítulo, nos permitirá probar si más de dos medias de población pueden considerarse iguales.

Inferencias acerca de varianzas de población

Es claro que no siempre estaremos interesados en medias y proporciones. Existen muchas situaciones administrativas en las que la preocupación se centrará en la variabilidad de una población. En la sección 11.5 se estudiará cómo utilizar la distribución ji-cuadrada para formar intervalos de confianza y probar hipótesis acerca de la varianza de una población. En la sección 11.6 veremos que las hipótesis que comparan las varianzas de dos poblaciones pueden probarse mediante la distribución F.

Ejercicios 11.1

- **11-1** ¿Por qué utilizamos una prueba ji-cuadrada?
- **11-2** ¿Por qué utilizamos el análisis de varianza?

■ **11-3** En cada una de las siguientes situaciones establezca si se debe hacer una prueba ji-cuadrada, análisis de varianza o inferencias acerca de la población.

a) Deseamos saber si la varianza de las temperaturas en primavera es la misma en la costa del Golfo que en la costa del Pacífico.
b) Queremos ver si la velocidad promedio de los automóviles que circulan por la carretera interestatal 95 cambia dependiendo del día de la semana.
c) Deseamos ver si el desempeño a largo plazo de las acciones en Wall Street (clasificado como bueno, promedio o pobre) es independiente del tamaño de la compañía (clasificada como pequeña, media o grande).
d) Antes de probar si $\mu_1 = \mu_2$, queremos saber si la suposición de que $\sigma_1^2 = \sigma_2^2$ es algo razonable.

■ **11-4** Responda si es verdadero o falso y explique sus respuestas.

a) Después de leer este capítulo, usted sabrá cómo hacer inferencias acerca de dos o más varianzas de población.
b) Después de leer este capítulo, usted sabrá cómo hacer inferencias acerca de dos o más medias de población.
c) Después de leer este capítulo, usted sabrá cómo hacer inferencias acerca de dos o más proporciones de población.

■ **11-5** Como ayuda para recordar qué distribución o técnica se utiliza, complete la siguiente tabla con el nombre de una distribución o de la técnica implicada. La clasificación de los renglones se refiere al número de parámetros involucrados en una prueba y la clasificación de las columnas al tipo de parámetro implicado. Algunas celdas pueden no tener entrada, otras pueden tener más de una entrada posible.

	Tipo de parámetro		
Número de parámetros implicados	μ	σ	***P***
1			
2			
3 o más			

11.2 Ji-cuadrada como prueba de independencia

Las diferencias muestrales entre proporciones, ¿son o no significativas?

En muchas ocasiones, los administradores necesitan saber si las diferencias que observan entre varias proporciones de la muestra son significativas o sólo se deben al azar. Suponga que el administrador de campaña de un candidato a la presidencia estudia tres regiones geográficas y encuentra que el 35, 42 y 51%, respectivamente, de los votantes investigados de las tres regiones reconocen el nombre del candidato. Si esta diferencia es significativa, el administrador puede llegar a la conclusión de que el lugar afectará la forma en que debe actuar el candidato. Pero si la diferencia no es significativa (es decir, si el administrador concluye que la diferencia solamente se debe al azar), entonces puede decidir que el lugar escogido para pronunciar un discurso proselitista no tendrá efecto sobre su recepción. Para conducir la campaña con éxito, entonces, el administrador necesita determinar si el lugar y el reconocimiento del nombre del candidato son dependientes o independientes.

Tablas de contingencia

Descripción de una tabla de contingencia

Suponga que en cuatro regiones, la Compañía Nacional de Cuidado de la Salud, de Estados Unidos, muestrea las actitudes de los empleados de sus hospitales respecto a la evaluación del desempeño en el trabajo. Los trabajadores eligen entre el método actual (dos evaluaciones al año) y un nuevo método propuesto (evaluaciones trimestrales). La tabla 11-1 se conoce como *tabla de contingencia* e ilustra la respuesta a esta pregunta que dio la muestra encuestada. Una tabla como ésta está formada por renglones y columnas: los renglones corren de manera horizontal y las columnas verticalmente. Note que las cuatro columnas de la tabla 11-1 proporcionan una base de clasificación —regiones

geográficas— y los dos renglones clasifican la información de otra manera: preferencia por los métodos de evaluación. La tabla 11-1 se conoce como tabla de contingencia 2×4, ya que consta de dos renglones y cuatro columnas. Describimos las dimensiones de una tabla de contingencia estableciendo primero el número de renglones y luego el número de columnas. La columna y el renglón con el "total" no cuentan como parte de las dimensiones.

Frecuencias observadas y esperadas

Planteamiento del problema en símbolos

Suponga que ahora simbolizamos las proporciones verdaderas de la población total de empleados que prefieren el plan actual como:

- p_N ← Proporción de empleados en el noreste que prefieren el plan actual
- p_S ← Proporción de empleados en el sureste que prefieren el plan actual
- p_C ← Proporción de empleados de la región central que prefieren el plan actual
- p_W ← Proporción de empleados de la región de la costa oeste que prefieren el plan actual

Utilizando estos símbolos, podemos establecer las hipótesis nula y alternativa de la siguiente manera:

$$H_0: p_N = p_S = p_C = p_W \leftarrow \text{Hipótesis nula}$$

$$H_1: p_N, p_S, p_C, p_W \text{ no son iguales} \leftarrow \text{Hipótesis alternativa}$$

Si la hipótesis nula es verdadera, podemos combinar los datos de las cuatro muestras y luego estimar la proporción de la fuerza de trabajo total (la población total) que prefiere el método de evaluación actual:

Proporción combinada que prefiere el método actual, suponiendo que la hipótesis nula de que no hay diferencia es verdadera

$$= \frac{68 + 75 + 57 + 79}{100 + 120 + 90 + 110}$$

$$= \frac{279}{420}$$

$$= 0.6643$$

Determinación de frecuencias esperadas

Obviamente, si el valor 0.6643 estima la proporción de población esperada que prefiere el método presente de evaluación, entonces 0.3357 ($= 1 - 0.6643$) es la estimación de la proporción esperada de la población que prefiere el método propuesto. Utilizando 0.6643 como *estimación* de la proporción de la población que prefiere el método de evaluación actual y 0.3357 como la *estimación* de la proporción de la población que prefiere el nuevo método, podemos estimar el número de empleados de la muestra de cada región que podríamos esperar que prefieran cada uno de los métodos de evaluación. La tabla 11-2 presenta estos cálculos.

Comparación de frecuencias esperada y observada

La tabla 11-3 combina toda la información contenida en las tablas 11-1 y 11-2. En ella se ilustran tanto la frecuencia real u observada, como la teórica o esperada, de trabajadores de la muestra que prefieren cada uno de los métodos de evaluación. Recuerde que las *frecuencias esperadas*, presentadas en negritas, fueron estimadas a partir de nuestra estimación combinada de la proporción.

Tabla 11-1

Respuesta de la muestra concerniente a los programas de evaluación de empleados en hospitales de la Compañía Nacional de Cuidado a la Salud

	Noreste	Sureste	Central	Costa oeste	Total
Número de empleados que prefieren el método actual	68	75	57	79	**279**
Número de empleados que prefieren el nuevo método	32	45	33	31	**141**
Total de empleados muestreados en cada región	**100**	**120**	**90**	**110**	**420**

Tabla 11-2

Proporción de empleados muestreados en cada región que se espera prefieran los dos métodos de evaluación

	Noreste	Sureste	Central	Costa oeste
Número total muestreado	100	120	90	110
Proporción estimada que prefiere el método actual	×0.6643	×0.6643	×0.6643	×0.6643
Número que se espera prefiera el método actual	**66.43**	**79.72**	**59.79**	**73.07**
Número total muestreado	100	120	90	110
Proporción estimada que prefiere el nuevo método	×0.3357	×0.3357	×0.3357	×0.3357
Número que se espera prefiera el nuevo método	**33.57**	**40.28**	**30.21**	**36.93**

Tabla 11-3

Comparación de las frecuencias observadas y esperadas de trabajadores muestreados

	Noreste	Sureste	Central	Costa occidental
FRECUENCIA CON QUE PREFIEREN EL MÉTODO ACTUAL:				
Frecuencia observada (real)	68	75	57	79
Frecuencia esperada (teórica)	**66.43**	**79.72**	**59.79**	**73.07**
FRECUENCIA CON QUE PREFIEREN EL NUEVO MÉTODO:				
Frecuencia observada (real)	32	45	33	31
Frecuencia esperada (teórica)	**33.57**	**40.28**	**30.21**	**36.93**

Razonamiento intuitivo acerca de las pruebas ji-cuadrada

Para probar la hipótesis nula, $p_N = p_S = p_C = p_W$, debemos comparar las frecuencias que se observaron (números en cursivas de la tabla 11-3) con las frecuencias que esperaríamos si la hipótesis nula fuera verdadera (números en negritas). Si los conjuntos de frecuencias observadas y esperadas son casi iguales, podemos razonar de manera intuitiva que la hipótesis nula se acepta. Si existe una diferencia grande entre estas frecuencias, podemos rechazar la hipótesis nula intuitivamente y concluir que existen diferencias significativas en las proporciones de empleados de las cuatro regiones que prefieren el nuevo método.

El estadístico ji-cuadrada

Para ir más allá de nuestra intuición acerca de las frecuencias observadas y esperadas, podemos usar el estadístico ji-cuadrada, que se calcula de la siguiente manera:

Cálculo del estadístico ji-cuadrada

Estadístico ji-cuadrada

Una frecuencia observada — Una frecuencia esperada

$$\chi^2 = \Sigma \frac{(f_o - f_e)^2}{f_e} \qquad [11\text{-}1]$$

Ji-cuadrada — Símbolo que significa "la suma de"

Esta fórmula establece que ji-cuadrada, o χ^2, es la suma que obtendremos si:

1. restamos f_e de f_o para cada uno de los ocho valores de la tabla 11-3;
2. elevamos al cuadrado cada diferencia;

Tabla 11-4

Cálculo del estadístico χ^2 (ji-cuadrada) a partir de los datos de la tabla 11-3

		Paso 1	Paso 2	Paso 3
f_o	f_e	$f_o - f_e$	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$
68	66.43	1.57	2.46	0.0370
75	79.72	−4.72	22.28	0.2795
57	59.79	−2.79	7.78	0.1301
79	73.07	5.93	35.16	0.4812
32	33.57	−1.57	2.46	0.0733
45	40.28	4.72	22.28	0.5531
33	30.21	2.79	7.78	0.2575
31	36.93	−5.93	35.16	0.9521
				2.7638

Paso 4 $\sum \frac{(f_o - f_e)^2}{f_e} = 2.764 \leftarrow \chi^2$ (ji-cuadrada)

3. dividimos cada diferencia al cuadrado entre f_e, y
4. sumamos los ocho resultados.

Numéricamente, los cálculos son sencillos si utilizamos una tabla como la 11-4, que muestra los pasos.

Interpretación del estadístico ji-cuadrada

La respuesta obtenida, 2.764, es el valor de ji-cuadrada en nuestro problema de comparación de las preferencias de métodos de evaluación. Si este valor fuera muy grande, digamos 20, indicaría una diferencia sustantiva entre los valores observados y los valores esperados. Una ji-cuadrada de cero, por otro lado, indica que las frecuencias observadas son exactamente iguales a las frecuencias esperadas. El valor de ji-cuadrada nunca puede ser negativo, porque la diferencia entre las frecuencias observadas y esperadas siempre están *al cuadrado*.

La distribución ji-cuadrada

Descripción de una distribución ji-cuadrada

Si la hipótesis nula es verdadera, entonces la distribución de muestreo del estadístico ji-cuadrada, χ^2, puede aproximarse bastante bien mediante una curva continua conocida como *distribución ji-cuadrada*. Como en el caso de la distribución *t*, existe una distribución ji-cuadrada diferente para cada número de grados de libertad. En la figura 11-1 se indican las tres diferentes distribuciones ji-cuadrada correspondientes a 1, 5 y 10 grados de libertad. Para un número muy pequeño de grados de libertad, la distribución ji-cuadrada está seriamente sesgada a la derecha. Conforme aumenta el número de grados de libertad, la curva hace cada vez más simétrica hasta que el número de grados de libertad alcanza valores grandes, en cuyo caso la distribución puede aproximarse con la normal.

Búsqueda de probabilidades cuando se utiliza una distribución ji-cuadrada

La distribución ji-cuadrada es una distribución de probabilidad. En consecuencia, el área total bajo la curva de cada distribución ji-cuadrada es 1.0. Como en el caso de la distribución *t*, es posible tener un número muy grande de distribuciones ji-cuadrada, de modo que no resulta práctico construir una tabla que contenga las áreas bajo la curva de todos los valores posibles del área. En la tabla 5 del apéndice se ilustran sólo las áreas de la cola que se utilizan con más frecuencia en pruebas de significancia que usan la distribución ji-cuadrada.

Determinación de los grados de libertad

Cálculo de los grados de libertad

Para utilizar la prueba ji-cuadrada, debemos calcular los grados de libertad en la tabla de contingencia aplicando la ecuación 11-2:

Grados de libertad en una prueba ji-cuadrada de independencia

$$\text{Número de grados de libertad} = (\text{número de renglones} - 1)(\text{número de columnas} - 1) \quad [11\text{-}2]$$

Examinemos qué tan apropiada es esta ecuación. Suponga que tenemos una tabla de contingencia de 3×4 como la figura 11-2. Conocemos los totales de renglón y columna, representados como RT_1, RT_2, RT_3 y CT_1, CT_2, CT_3 y CT_4. Como lo analizamos en el capítulo 7, el número de grados de libertad es igual al número de valores que podemos especificar libremente.

Observe ahora el primer renglón de la tabla de contingencia de la figura 11-2. Una vez especificados los primeros tres valores de ese renglón (denotados por palomitas), el cuarto valor en ese mismo renglón (denotado por un círculo) ya está determinado; no tenemos la libertad de especificarlo, debido a que conocemos el total del renglón.

Del mismo modo, en el segundo renglón de la tabla de contingencia de la figura 11-2, cuando especificamos los tres primeros valores (otra vez palomitas), el cuarto valor queda determinado y no lo podemos especificar libremente. Este cuarto valor se denota por un círculo.

Pasando ahora al tercer renglón, nos damos cuenta de que su primera entrada está determinado, porque *ya conocemos las primeras dos entradas de la primera columna y su total*; de nuevo, esta entrada tiene un círculo. Podemos aplicar este mismo razonamiento a la segunda y tercera entradas del tercer renglón, que también tienen círculos.

Por último, en la última entrada del tercer renglón (denotada por un asterisco), nos damos cuenta de que no podemos especificar libremente su valor, porque ya se conocen las dos primeras entradas de la cuarta columna. Contando el número de palomitas de la tabla de contingencias de la figura 11-2, se puede ver que el número de valores que podemos especificar libremente es 6. Esto es igual a 2×3, o (número de renglones $-$ 1) (número de columnas $-$ l).

Esto es justo lo que tenemos en la ecuación 11-2. En la tabla 11-5 se ilustran las dimensiones de renglones y columnas de tres tablas de contingencia más y se indican los grados de libertad apropiados en cada caso.

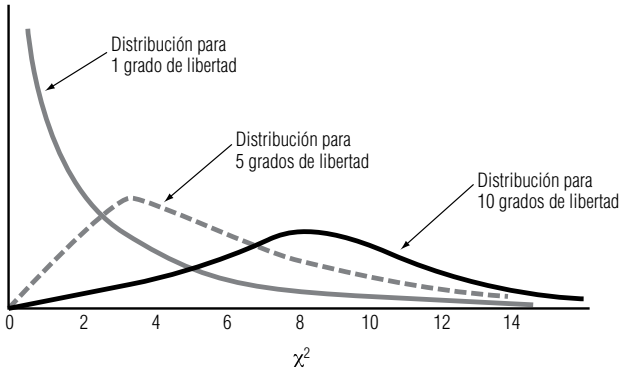


FIGURA 11-1

Distribuciones ji-cuadradas con 1, 5 y 10 grados de libertad

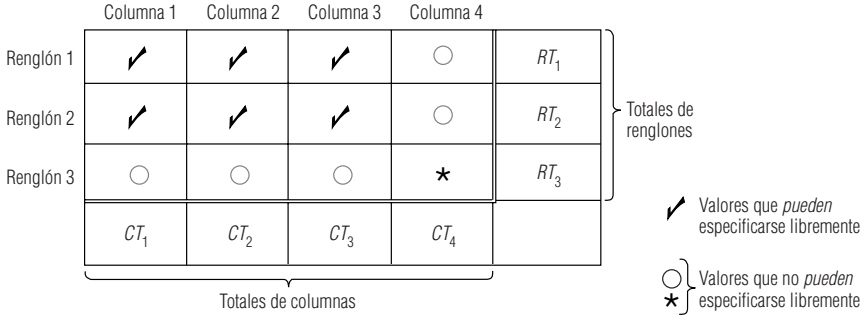


FIGURA 11-2

Tabla de contingencia 3×4 para determinar el número de grados de libertad

Uso de la prueba ji-cuadrada

Planteamiento del problema en símbolos

Regresando al ejemplo de las preferencias de evaluación del trabajo en los hospitales de la Compañía Nacional de Cuidado de la Salud, utilizamos la prueba ji-cuadrada para determinar si la actitud hacia los procedimientos de evaluación es independiente de la región geográfica. Si la compañía desea probar la hipótesis nula a un nivel de significancia de 0.10, nuestro problema puede resumirse de la siguiente manera:

H_0: $p_N = p_S = p_C = p_W$ ← Hipótesis nula

H_1: p_N, p_S, p_C y p_W no son iguales ← Hipótesis alternativa

$\alpha = 0.10$ ← Nivel de significancia para probar esta hipótesis

Cálculo de los grados de libertad

Como la tabla de contingencia para este problema (tabla 11-1) tiene dos renglones y cuatro columnas, el número adecuado de grados de libertad es:

Número de grados de libertad

$$(r - 1)(c - 1) \qquad [11\text{-}2]$$

Número de renglones — Número de columnas

$$= (2 - 1)(4 - 1)$$
$$= (1)(3)$$
$$= 3 \leftarrow \text{Grados de libertad}$$

Ilustración de la prueba de hipótesis

La figura 11-3 ilustra una distribución ji-cuadrada con tres grados de libertad (el nivel de significancia es el área sombreada). En la tabla 5 del apéndice podemos buscar en la columna 0.10, hacia abajo, hasta el renglón de 3 grados de libertad. Ahí se encuentra el valor del estadístico ji-cuadrada, 6.251. Podemos interpretar esto como que con 3 grados de libertad, la región a la derecha del valor ji-cuadrada, 6.251, contiene 0.10 del área bajo la curva. Entonces, la región de aceptación de la hipótesis nula en la figura 11-3 va de la cola izquierda de la curva al valor ji-cuadrada de 6.251.

Interpretación de los resultados

Como podemos ver de la figura 11-3, el valor ji-cuadrada de la muestra, 2.764, calculado en la tabla 11-4, cae dentro de la región de aceptación. Por tanto, aceptamos la hipótesis nula de que no existe diferencia entre las actitudes con respecto a la evaluación del trabajo en las cuatro regiones geográficas. En otras palabras, concluimos que la actitud hacia la evaluación del desempeño es independiente de la región.

Tablas de contingencia con más de dos renglones

¿Son independientes la permanencia en un hospital y la cobertura de seguro?

El señor George McMahon, presidente de la Compañía Nacional General de Seguros de Salud, se opone al seguro nacional de salud. Argumenta que su implementación sería muy costosa, en particular debido a que la existencia de este sistema tendería a fomentar permanencias hospitalarias más prolongadas, además de otros efectos. George piensa que el tiempo de hospitalización depende del tipo de seguro de salud que tengan las personas. Le pide a Donna McClish, la especialista en esta-

Tabla 11-5

Determinación de los grados de libertad en tres tablas de contingencia

Tabla de contingencia	Número de renglones (r)	Número de columnas (c)	$r - 1$	$c - 1$	Grados de libertad $(r - 1)(c - 1)$
A	3	4	$3 - 1 = 2$	$4 - 1 = 3$	$(2)(3) = 6$
B	5	7	$5 - 1 = 4$	$7 - 1 = 6$	$(4)(6) = 24$
C	6	9	$6 - 1 = 5$	$9 - 1 = 8$	$(5)(8) = 40$

Tabla 11-6

Datos de hospitalizaciones clasificados según el tipo de cobertura del seguro y el tiempo de estancia

		Días en el hospital			
		<5	5–10	>10	Total
Fracción de costos cubiertos por el seguro	<25%	40	75	65	**180**
	25–50%	30	45	75	**150**
	>50%	40	100	190	**330**
	Total	**110**	**220**	**330**	**660**

dística de la empresa, que verifique el asunto. Donna colectó datos de una muestra aleatoria de 660 hospitalizaciones y los resumió en la tabla 11-6.

Esta tabla da las frecuencias observadas para nueve hospitalizaciones diferentes y el tipo de seguro (o "celdas") en que dividimos la muestra. Donna desea probar las hipótesis:

Planteamiento de las hipótesis

H_0: el tiempo de estancia y tipo de seguro son independientes

H_1: el tiempo de estancia depende del tipo de seguro

$\alpha = 0.01$ ← Nivel de significancia para probar estas hipótesis

Búsqueda de frecuencias esperadas

Utilizaremos una prueba ji-cuadrada, de manera que primero tenemos que hallar las frecuencias esperadas para cada una de las nueve celdas. Veamos cómo podemos encontrarlas observando la celda que corresponde a estancias de menos de 5 días y una cobertura de seguro de menos del 25% de los costos.

Estimación de las proporciones correspondientes a las celdas

Un total de 180 de las 660 estancias registradas en la tabla 11-6 tienen una cobertura de seguro de menos del 25% de los costos. Entonces la cifra 180/660 *estima* la proporción de la población que tiene una cobertura menor que el 25% de los costos. Similarmente, 110/660 *estima* la proporción de todas las hospitalizaciones con duración menor que 5 días. Si la duración de la estancia y el tipo de seguro son en realidad independientes, podemos utilizar la ecuación 4-4 para *estimar* la proporción en la primera celda (menos de 5 días y cobertura meno que el 25% de los costos).

Definimos:

- A = el evento "una estancia corresponde a alguien cuyo seguro cubre menos del 25% de los costos"
- B = el evento "una estancia menor que 5 días"

Entonces,

$$\begin{aligned} P(\text{primera celda}) &= P(A \text{ y } B) \\ &= P(A) \times P(B) \\ &= \left(\frac{180}{660}\right)\left(\frac{110}{660}\right) \\ &= 1/22 \end{aligned} \qquad [4\text{-}4]$$

Como 1/22 es la *proporción* esperada en la primera celda, la *frecuencia* esperada en esa celda es:

$$(1/22)(660) = 30 \text{ observaciones}$$

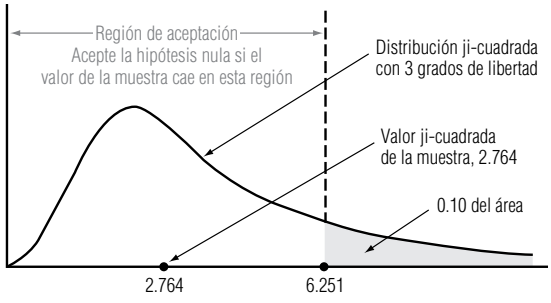


FIGURA 11-3

Prueba de hipótesis ji-cuadrada al nivel de significancia de 0.10 que ilustra la región de aceptación y el valor ji-cuadrada de la muestra de 2.764

En general, podemos calcular la frecuencia esperada para cualquier celda con la ecuación 11-3:

Cálculo de las frecuencias esperadas para las celdas

Frecuencia esperada para cualquier celda

$$f_e = \frac{RT \times CT}{n} \qquad [11\text{-}3]$$

donde,

- f_e = frecuencia esperada en una celda dada
- RT = total por renglón para el renglón que contiene a esa celda
- CT = total por columna para la columna que contiene a esa celda
- n = número total de observaciones

Ahora podemos utilizar las ecuaciones 11-3 y 11-1 para calcular todas las frecuencias esperadas y el valor del estadístico ji-cuadrada. Los cálculos se hicieron en la tabla 11-7.

La figura 11-4 ilustra una distribución ji-cuadrada con cuatro grados de libertad (número de renglones $-$ 1 = 2) $\times$ (número de columnas $-$ 1 = 2), donde el nivel de significancia 0.01 es el área sombreada. La tabla 5 del apéndice (en la columna 0.01 y el renglón de 4 grados de libertad) indica a Donna que para su problema, la región a la derecha de un valor ji-cuadrada 13.277 contiene 0.01 del área bajo la curva. Por consiguiente, la región de aceptación para la hipótesis nula en la figura 11-4 va desde la cola izquierda de la curva hasta el valor ji-cuadrada, 13.277.

Interpretación de los resultados de la prueba

Como se ve en la figura 11-4, el valor ji-cuadrada de la muestra, 24.315, que Donna calculó en la tabla 11-7, no se encuentra dentro de la región de aceptación. Así, ella debe rechazar la hipótesis nula e informar al señor McMahon que la evidencia refuerza su creencia de que la duración de las hospitalizaciones y la cobertura de los seguros son dependientes entre sí.

Tabla 11-7

Cálculo de las frecuencias esperadas y ji-cuadrada a partir de los datos de la tabla 11-6

Renglón	Columna	f_o	$f_e = \frac{RT \times CT}{n}$	$f_o - f_e$	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$
1	1	40	$\frac{180 \times 110}{660}$ = 30	10	100	3.333
1	2	75	$\frac{180 \times 220}{660}$ = 60	15	225	3.750
1	3	65	$\frac{180 \times 330}{660}$ = 90	−25	625	6.944
2	1	30	$\frac{150 \times 110}{660}$ = 25	5	25	1.000
2	2	45	$\frac{150 \times 220}{660}$ = 50	−5	25	0.500
2	3	75	$\frac{150 \times 330}{660}$ = 75	0	0	0.000
3	1	40	$\frac{330 \times 110}{660}$ = 55	−15	225	4.091
3	2	100	$\frac{330 \times 220}{660}$ = 110	−10	100	0.909
3	3	190	$\frac{330 \times 330}{660}$ = 165	25	625	3.788

[11-1] $\sum \frac{(f_o - f_e)^2}{f_e} = 24.315 \leftarrow \chi^2$ (ji-cuadrada)

Precauciones al usar la prueba ji-cuadrada

Use tamaños de muestra grandes

Para utilizar una prueba de hipótesis ji-cuadrada, debemos tener un tamaño de muestra lo suficientemente grande que garantice la similitud entre la distribución teórica correcta y nuestra distribución de muestreo de χ^2, el estadístico ji-cuadrada. Cuando las frecuencias esperadas son muy pequeñas, el valor de χ^2 estará sobrestimado y se tendrá como resultado demasiados rechazos de la hipótesis nula. **Para evitar incurrir en inferencias incorrectas de la prueba de hipótesis ji-cuadrada, siga la regla general de que una frecuencia esperada de menos de 5 en una celda de una tabla de contingencia es demasiado pequeña para utilizarse.*** Cuando la tabla contiene más de una celda con una frecuencia esperada menor que 5, podemos combinarlas con el fin de obtener una frecuencia esperada de 5 o más grande. Sin embargo, al hacerlo, reducimos el número de categorías de los datos y obtendremos menos información de la tabla de contingencia.

Utilice con cuidado los datos recolectados

Esta regla nos permitirá utilizar la prueba de hipótesis ji-cuadrada de una manera apropiada, pero, desafortunadamente, cada prueba sólo puede reflejar (y no mejorar) la calidad de los datos con que se hace. Hasta este momento, hemos rechazado la hipótesis nula si la diferencia entre las frecuencias observadas y esperadas —es decir, el valor ji-cuadrada calculado— es demasiado grande. En el caso de la preferencia por el método de evaluación, habríamos rechazado la hipótesis nula a un nivel de significancia de 0.10 si nuestro valor ji-cuadrada hubiera sido 6.251 o más. **Pero si el valor ji-cuadrada fuera cero, tendríamos que ser cuidadosos al preguntar si no existe *absolutamente ninguna diferencia* entre las frecuencias observadas y las esperadas.** Si tenemos una opinión fuerte de que *debería* existir alguna diferencia, tendríamos que examinar tanto la forma de recolección de los datos o como la manera de hacer las mediciones, o ambas cosas, para tener la certeza de que las diferencias existentes no fueron minimizadas o pasadas por alto al recolectar los datos de muestra.

Datos de los chícharos de Mendel

En la década de 1860, sus experimentos con las características de los chícharos condujeron al monje Gregor Mendel a proponer la existencia de los genes. Los resultados experimentales de Mendel eran sorprendentemente cercanos a los pronosticados por su teoría. Un siglo más tarde, los estadísticos estudiaron los "datos de los chícharos" de Mendel, les hicieron una prueba ji-cuadrada y llegaron a la conclusión de que el valor ji-cuadrada era demasiado pequeño; es decir, los datos experimentales dados por Mendel eran tan cercanos a los valores esperados que sólo pudieron concluir que había manipulado los datos.

Uso de la computadora para realizar pruebas ji-cuadrada

Uso de SPSS para una prueba ji-cuadrada

A pesar de que los cálculos necesarios para una prueba ji-cuadrada de independencia son relativamente sencillos, procesar grandes conjuntos de datos se puede convertir en algo tedioso. Casi todos los paquetes de computación estadísticos que más se usan contienen rutinas para realizar estas pruebas. En la figura 11-5 se observan los resultados que obtuvimos cuando utilizamos el paquete SPSS para analizar los datos de las hospitalizaciones dados en la tabla 11-6. Comparemos los resultados del paquete con el análisis que hicimos a mano en páginas anteriores.

Comparación de los resultados de computadora con los calculados a mano

En cada celda de la figura 11-5, SPSS imprime la frecuencia observada (f_o), la frecuencia esperada (f_e) y la contribución de esa celda al valor del estadístico χ^2 [$f_o - f_e)^2/f_e$]. Luego, en la parte inferior de la tabla, SPSS imprime el valor ji-cuadrada de la muestra, el número de grados de libertad y un valor *p*, donde el valor *p* es la probabilidad de obtener un valor ji-cuadrada observado igual (o mayor) que el valor ji-cuadrada de la muestra, si la hipótesis de independencia es válida.

Interpretación de los resultados

Recordando el análisis de los valores *p* del capítulo 9, sabemos que rechazaremos H_0 si el valor *p* es menor que α, el nivel de significancia de la prueba. En nuestro ejemplo, $\alpha = 0.01$ y el valor *p* calculado por SPSS es 0.0001, de modo que otra vez rechazamos H_0 y concluimos que la duración de la estancia y la cobertura del seguro no son independientes.

*Los especialistas en estadística han desarrollado factores de corrección que, en algunos casos, permiten utilizar celdas con frecuencias esperadas menores que 5. La derivación y uso de tales factores de corrección están más allá de los objetivos de este libro.

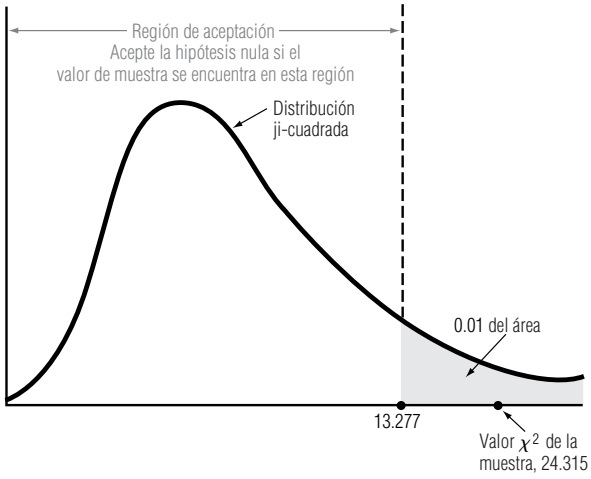


FIGURA 11-4

Prueba de hipótesis ji-cuadrada al nivel de significancia de 0.01 que ilustra la región de aceptación y el valor ji-cuadrada de la muestra de 24.315

EJEMPLO DEL USO DE SPSS PARA UNA PRUEBA DE INDEPENDENCIA

DÍAS = DURACIÓN DE LA ESTANCIA
COBERTURA = % DE COSTOS CUBIERTOS POR EL SEGURO

TABLA DE DÍAS POR COBERTURA

DÍAS COBERTURA

FRECUENCIA ESPERADA CELDA JI2	MENOS DE 25%	25 - 50%	MÁS DE 50%	TOTAL
MENOS DE 5	40	30	40	110
	30.0	25.0	55.0	
	3.33333	1	4.09091	
5 A 10	75	45	100	220
	60.0	50.0	110.0	
	3.75	0.5	.909091	
MÁS DE 10	65	75	190	330
	90.0	75.0	165.0	
	6.94444	0	3.78788	
TOTAL	180	150	330	660

ESTADÍSTICOS PARA LA TABLA DE DÍAS POR COBERTURA

ESTADÍSTICO	GL	VALOR	PROB
JI-CUADRADA	4	24.316	0.0001

FIGURA 11-5

Resultados obtenidos con SPSS para el problema de las hospitalizaciones

Uso de Minitab para una prueba ji-cuadrada

Como un segundo ejemplo basado en la computadora para la prueba ji-cuadrada, regresemos a los datos de ingresos de la tabla 11 del apéndice. Se usó Minitab para codificar el cambio de un año a otro en los ingresos del último trimestre en cinco grupos:

Cambio	Grupo
$x < -\$0.20$	−2
$-\$0.20 \le x < -\$0.05$	−1
$-\$0.05 \le x < \$0.05$	0
$\$0.05 \le x < \$0.20$	1
$\$0.20 \le x$	2

Además, la variable CAMBIO se convirtió de su valor de caracteres O, A y N en valores numéricos 1, 2 y 3 en una nueva variable llamada MERCADO.

Estadísticas tabuladas

RENGLONES:	GRUPO	COLUMNAS:	MERCADO	
	1	2	3	TODO
-2	24	8	17	49
-1	14	7	10	31
0	41	11	11	63
1	23	6	17	46
2	9	6	20	35
TODO	111	38	75	224

JI-CUADRADA = 19.101 CON G.L. = 8

CONTENIDO DE CELDAS --
CUENTA

FIGURA 11-6

Uso de Minitab para realizar una prueba de independencia

Ahora, probemos, para $\alpha = 0.01$, si el cambio año con año en los ingresos del último trimestre del año depende de dónde se negociaron los valores de la compañía.

Planteamiento de las hipótesis

H_0: el cambio en los ingresos y la bolsa de valores son independientes

H_1: el cambio en los ingresos depende de la bolsa de valores

$\alpha = 0.01$ ← Nivel de significancia para probar estas hipótesis

Interpretación de resultados

En la figura 11-6 se usó Minitab para realizar esta prueba de independencia. En cada celda de la figura, Minitab imprime las frecuencias observadas (f_o); también proporciona el total de renglones y columnas. En la parte inferior, da el valor de ji-cuadrada de la muestra y los grados de libertad. Para $\alpha = 0.01$ con 8 grados de libertad, el valor crítico ji-cuadrada es 20.090. Por tanto, no se puede rechazar H_0. Se puede concluir que los cambios en los ingresos y la bolsa de valores que se elige son independientes.

SUGERENCIAS Y SUPOSICIONES

Advertencia: los renglones y columnas de una tabla de contingencia ji-cuadrada *deben* ser categorías mutuamente excluyentes y exhaustivas, es decir, que agoten *todas* las posibilidades de la muestra. Sigerencia: piense que las celdas son como pequeñas cajas y cada elemento de la muestra es una canica. Cada canica debe colocarse en una caja y no pueden quedar canicas fuera si se quiere que la prueba sea válida. Por ejemplo, un sondeo de electores que tiene celdas en la tabla de contingencia sólo para demócratas y republicanos ignora las opiniones de los electores no afiliados. Sugerencia: piense en las categorías "dueño de auto" y "dueño de bicicleta" que no permiten que las personas posean ambos.

Ejercicios 11.2

Ejercicios de autoevaluación

EA 11-1 Un gerente de marca está preocupado porque la participación de mercado de su marca se distribuye en forma dispareja en el país. En una encuesta en la que se dividió al país en cuatro regiones geográficas, se tomó un muestreo aleatorio de 100 consumidores en cada región, con los siguientes resultados:

	REGIÓN				
	NE	NO	SE	SO	TOTAL
Compra la marca	40	55	45	50	190
No compra la marca	60	45	55	50	210
Total	**100**	**100**	**100**	**100**	**400**

Desarrolle una tabla de frecuencias observadas y esperadas para este problema.

EA 11-2 Para el ejercicio EV 11-1:

a) Calcule el valor χ^2 de la muestra.
b) Establezca las hipótesis nula y alternativa.
c) Para $\alpha = 0.05$, pruebe si la participación de la marca es la misma en las cuatro regiones.

Conceptos básicos

11-6 Para cada una de las dimensiones siguientes de tablas de contingencia, ¿cuántos grados de libertad tendrá el estadístico ji-cuadrada?

a) 5 renglones, 4 columnas.
b) 6 renglones, 2 columnas.
c) 3 renglones, 7 columnas.
d) 4 renglones, 4 columnas.

Aplicaciones

11-7 Una agencia de publicidad intenta determinar la composición demográfica del mercado para un nuevo producto. Seleccionaron al azar 75 personas de cada uno de 5 grupos de edad y les presentaron el producto. Los resultados de la encuesta son los siguientes:

	Grupo de edad				
Actividad futura	**18-29**	**30-39**	**40-49**	**50-59**	**60-69**
Compra frecuente	12	18	17	22	32
Compra alguna vez	18	25	29	24	30
Nunca compra	45	32	29	29	13

Desarrolle una tabla de frecuencias observadas y esperadas para este problema.

11-8 Para el ejercicio 11-7:

a) Calcule el valor χ^2 de la muestra.
b) Establezca las hipótesis nula y alternativa.
c) Si el nivel de significancia es 0.01, ¿debe rechazarse la hipótesis nula?

11-9 Para ver si las ventas de chips de silicio son independientes del punto del ciclo de negocios en que se encuentre la economía de Estados Unidos se han recogido datos de las ventas semanales de Zippy Chippy, una empresa de Silicon Valley, y datos acerca de si la economía de Estados Unidos subía al pico del ciclo, estaba en el pico, iba a la baja o estaba en el punto bajo. Los resultados son los siguientes:

	VENTA SEMANAL DE CHIPS			
	Alta	**Mediana**	**Baja**	**TOTAL**
Economía				
En el pico	20	7	3	30
En el punto bajo	30	40	30	100
Subiendo	20	8	2	30
Bajando	30	5	5	40
Total	**100**	**60**	**40**	**200**

Calcule una tabla de frecuencias observadas y esperadas para este problema.

11-10 Para el ejercicio 11-9:

a) Establezca las hipótesis nula y alternativa.
b) Calcule el valor χ^2 de la muestra.
c) Al nivel de significancia de 0.10, ¿cuál es su conclusión?

11-11 Un asesor financiero está interesado en las diferencias de estructura de capital respecto a compañías de distintos tamaños dentro de cierta industria. El asesor investiga un grupo de empresas con activos de diferentes cantidades y las organiza en tres grupos. Clasifica cada compañía según si su débito total es mayor que la cantidad de acciones ordinarias de los accionistas o si es menor que éstas. Los resultados de la investigación son:

	Tamaño del activo de la compañía (en miles de dólares)			
	<500	500-2,000	2,000+	Total
Deuda menor que cantidad de acciones	7	10	8	**25**
Deuda mayor que cantidad de acciones	10	18	9	**37**
Total	**17**	**28**	**17**	**62**

¿Los tres tamaños de empresas tienen la misma estructura de capital? Use un nivel de significancia de 0.10.

11-12 Un editor de periódicos que trata de determinar con precisión las características de su mercado, se pregunta si la costumbre de leer diarios en la comunidad se relaciona con el nivel educativo de las personas. Pregunta a los adultos del área acerca de su nivel educativo y a la frecuencia con que leen el periódico. Los resultados se muestran en la siguiente tabla:

	NIVEL EDUCATIVO				
FRECUENCIA CON LA QUE LEE	Profesional o posgrado	Pasante de licenciatura	Bachillerato	Bachillerato inconcluso	TOTAL
Nunca	10	17	11	21	59
Algunas veces	12	23	8	5	48
Mañana o tarde	35	38	16	7	96
Ambas ediciones	28	19	6	13	66
Total	**85**	**97**	**41**	**46**	**269**

A un nivel de significancia de 0.10, ¿la frecuencia con que leen el periódico en la comunidad difiere con el nivel de educación de los lectores?

11-13 Un educador opina que las calificaciones obtenidas por los alumnos de bachillerato dependen del tiempo que pasan escuchando música. Para probar esta teoría, ha repartido al azar 400 cuestionarios entre estudiantes. En ellos hay dos preguntas: ¿cuántas horas por semana escuchas música?, ¿qué promedio general de calificaciones tienes? Los datos de la encuesta se presentan en la tabla siguiente. Utilizando un nivel de significancia del 5%, pruebe si las calificaciones y el tiempo dedicado a escuchar música son independientes o dependientes.

HORAS CONSUMIDAS ESCUCHANDO MÚSICA	PROMEDIO GENERAL DE CALIFICACIONES					
	A	B	C	D	F	TOTAL
<5 h	13	10	11	16	5	55
5-10 h	20	27	27	19	2	95
11-20 h	9	27	71	16	32	155
> 20 h	8	11	41	24	11	95
Total	**50**	**75**	**150**	**75**	**50**	**400**

Soluciones a los ejercicios de autoevaluación

EA 11-1

	Región			
	NE	NO	SE	SO
Compradores				
Observada	40	55	45	50
Esperada	47.5	47.5	47.5	47.5
No compradores				
Observada	60	45	55	50
Esperada	52.5	52.5	52.5	52.5

EA 11-2 a)

f_o	f_e	$f_o - f_e$	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_o}$
40	47.5	−7.5	56.25	1.184
55	47.5	7.5	56.25	1.184
45	47.5	−2.5	6.25	0.132
50	47.5	2.5	6.25	0.132
60	52.5	7.5	56.25	1.071
45	52.5	−7.5	56.25	1.071
55	52.5	2.5	6.25	0.119
50	52.5	−2.5	6.25	0.119

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e} = 5.012$$

b) De dos formas, cualquiera de las dos es aceptable:
(1) H_0: la región es independiente de las compras
H_1: la región se relaciona con las compras (dependiente)
(2) H_0: $p_{ne} = p_{no} = p_{se} = p_{no}$
H_1: no todas las proporciones son iguales

c) Con $1 \times 3 = 3$ grados de libertad y $\alpha = 0.05$, el valor crítico de χ^2 es 7.815, por lo que no se rechaza H_0, dado que $5.012 < 7.815$. La participación de la marca no difiere de manera significativa con la región.

11.3 Ji-cuadrada como prueba de bondad de ajuste: prueba de lo apropiado de una distribución

En la sección anterior, utilizamos la prueba ji-cuadrada para decidir si aceptamos o no una hipótesis nula que era una hipótesis de independencia entre dos variables. En el ejemplo, estas dos variables eran la actitud hacia la evaluación del desempeño en el trabajo y la región geográfica.

Función de una prueba de bondad de ajuste

La prueba ji-cuadrada puede utilizarse también para decidir si una distribución de probabilidad en particular, como la binomial, la de Poisson o la normal, es la *apropiada*. Ésta es una habilidad importante, porque como tomadores de decisiones que utilizamos la estadística, necesitaremos escoger cierta distribución de probabilidad para representar la distribución de los datos que tengamos que analizar. Necesitaremos la habilidad para cuestionar hasta dónde podemos llegar con las suposición de que podemos usar una distribución en particular antes de que debamos concluir que esa distribución ya no se puede aplicar. La **prueba ji-cuadrada nos permite hacernos esta pregunta y probar si existe una diferencia significativa entre una distribución de frecuencias observada y una distribución de frecuencias teórica.** De esta manera, podemos determinar la *bondad de ajuste* de una distribución teórica (es decir, qué tan bien se ajusta a la distribución de los datos que observamos). De esta forma, podemos determinar si debemos creer que los datos observados constituyen una muestra obtenida de la distribución teórica hipotética.

Cálculo de las frecuencias observadas y esperadas

Suponga que la compañía Gordon requiere que los estudiantes del último año de la universidad que buscan trabajo sean entrevistados por tres ejecutivos diferentes. Esto permite a la compañía obtener una evaluación por consenso de los candidatos. Cada ejecutivo califica al candidato como positivo o negativo. La tabla 11-8 contiene los resultados de las entrevistas de los últimos 100 candidatos.

Con el propósito de planear la contratación, el director de selección de personal de la compañía piensa que el proceso de entrevistas puede ser aproximado por una distribución binomial con $p = 0.40$, es decir, con una posibilidad del 40% de que cualquier candidato obtenga una calificación positiva

Tabla 11-8
Resultados de las entrevistas de 100 candidatos

Calificaciones positivas posibles en las tres entrevistas	Núm. de candidatos que obtienen cada calificación
0	18
1	47
2	24
3	11
	100

en cualquiera de las entrevistas. Si el director desea probar esta hipótesis a un nivel de significancia de 0.20, ¿cómo debe proceder?

Planteamiento del problema en símbolos

H_O: una distribución binomial con $p = 0.40$ es una buena descripción del proceso de entrevistas ← Hipótesis nula

H_1: una distribución binomial con $p = 0.40$ *no* es una buena descripción del proceso de entrevistas ← Hipótesis alternativa

$\alpha = 0.20$ ← Nivel de significancia para la prueba de estas hipótesis

Cálculo de las probabilidades binomiales

Para resolver este problema, debemos determinar si las discrepancias entre las frecuencias observadas y las que esperaríamos (si la distribución binomial *es* el modelo apropiado) se pueden atribuir al azar. Podemos empezar por determinar cuáles serían las probabilidades binomiales para esta situación. Para las tres entrevistas, encontramos la probabilidad de éxito en la tabla de distribución binomial (tabla 3 del apéndice) buscando en la columna de $n = 3$ y $p = 0.40$. Los resultados se resumen en la tabla 11-9.

Ahora podemos utilizar las probabilidades binomiales teóricas de los resultados para calcular las frecuencias esperadas. Al comparar esas frecuencias esperadas con nuestras frecuencias observadas usando la prueba χ^2, podemos examinar la magnitud de la diferencia entre ellas. En la tabla 11-10 se dan las frecuencias observadas, las probabilidades binomiales de la tabla 11-9 y las frecuencias esperadas para la muestra de 100 entrevistas.

Tabla 11-9
Probabilidades binomiales para el problema de las entrevistas

Calificaciones positivas posibles en las tres entrevistas	Probabilidades binomiales para esos resultados
0	0.2160
1	0.4320
2	0.2880
3	0.0640
	1.0000

Tabla 11-10
Frecuencias observadas, probabilidades binomiales adecuadas y frecuencias esperadas para el problema de las entrevistas

Calificaciones positivas posibles en las tres entrevistas	Frecuencia observada de candidatos que obtienen estas calificaciones	Probabilidad binomial de resultados posibles		Número de candidatos entrevistados		Frecuencia esperada de candidatos que obtienen estas calificaciones
0	18	0.2160	×	100	=	21.6
1	47	0.4320	×	100	=	43.2
2	24	0.2880	×	100	=	28.8
3	11	0.0640	×	100	=	6.4
	100	**1.0000**				**100.0**

Tabla 11-11

Cálculo del estadístico χ^2 a partir de los datos de las entrevistas dados en la tabla 11-10

Frecuencia observada (f_o)	Frecuencia esperada (f_e)	$f_o - f_e$	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$
18	21.6	−3.6	12.96	0.6000
47	43.2	3.8	14.44	0.3343
24	28.8	−4.8	23.04	0.8000
11	6.4	4.6	21.16	3.3063
				5.0406

$$\Sigma \frac{(f_o - f_e)^2}{f_e} = 5.0406 \leftarrow \chi^2$$

Cálculo del estadístico ji-cuadrada

Para calcular el estadístico ji-cuadrada de este problema, podemos utilizar la ecuación 11-1:

$$\chi^2 = \Sigma \frac{(f_o - f_e)^2}{f_e} \qquad [11\text{-}1]$$

y el formato introducido en la tabla 11-4. Este proceso se ilustra en la tabla 11-11.

Determinación de los grados de libertad de una prueba de bondad de ajuste

Primero, cuente el número de clases

Antes de calcular el número adecuado de grados de libertad para una prueba ji-cuadrada de bondad de ajuste, es necesario contar el número de clases (denotado por k) para las que se compararon las frecuencias observadas y esperadas. El problema de las entrevistas contiene cuatro clases: 0, 1, 2 y 3 calificaciones positivas. Así, empezamos con cuatro grados de libertad. Sin embargo, como las cuatro frecuencias observadas deben sumar 100, el número total de frecuencias observadas que podemos especificar libremente es sólo $k - 1 = 3$. La cuarta queda determinada por la suma total de 100.

Luego reste los grados de libertad perdidos al estimar parámetros de población

Para resolver un problema de bondad de ajuste, podemos vernos forzados a imponer restricciones adicionales en el cálculo de los grados de libertad. Suponga que estamos utilizando la prueba ji-cuadrada como una prueba de bondad de ajuste para determinar si una distribución normal se ajusta a un conjunto de frecuencias observadas. Si tenemos seis clases de frecuencias observadas ($k = 6$), concluimos que solamente tenemos $k - 1 = 5$ grados de libertad. Sin embargo, si también tenemos que utilizar la media de la muestra como una estimación de la media de la población, tendremos que restar un grado de libertad adicional, lo cual nos deja con sólo 4. Después, si tenemos que utilizar la desviación estándar de la muestra para estimar la desviación estándar de la población, tendremos que restar un grado de libertad más, lo que deja 3. Nuestra regla general en estos casos es: **primero aplique la regla ($k - 1$) y luego reste un grado de libertad adicional por cada parámetro de población que debe estimar a partir de los datos de la muestra.**

En el ejemplo de las entrevistas, tenemos cuatro clases de frecuencias observadas. Como resultado, $k = 4$ y el número apropiado de grados de libertad es $k - 1 = 3$. No necesitarnos estimar ningún parámetro de población, así que no necesitamos reducir más este número.

Uso de la prueba ji-cuadrada de bondad de ajuste

Búsqueda del límite de la región de aceptación

En el problema de las entrevistas, la compañía desea probar la hipótesis de bondad de ajuste al nivel de significancia de 0.20. En la tabla 5 del apéndice, entonces, debemos buscar en la columna correspondiente a 0.20 el renglón de 3 grados de libertad. Tenemos que el valor del estadístico ji-cuadrada es 4.642. Podemos interpretar este valor de la siguiente manera: con 3 grados de libertad, la región que se encuentra a la derecha del valor ji-cuadrada, 4.642, contiene 0.20 del área bajo la curva.

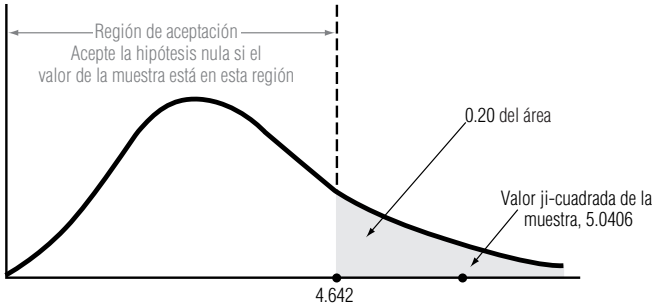


FIGURA 11-7

Prueba de bondad de ajuste al nivel de significancia de 0.20; ilustra la región de aceptación y el valor ji-cuadrada de la muestra, 5.0406

Ilustración del problema

Interpretación de los resultados

La figura 11-7 ilustra una distribución ji-cuadrada con 3 grados de libertad, en la que el nivel de significancia de 0.20 es el área sombreado. Note que la región de aceptación para la hipótesis nula (que establece que los datos de la muestra vienen de una distribución binomial con $p = 0.4$) se extiende desde la cola izquierda hasta el valor ji-cuadrada de 4.642. Obviamente, el valor ji-cuadrada de la muestra, 5.0406, cae fuera de esta región de aceptación. Por consiguiente, rechazamos la hipótesis nula y llegamos a la conclusión de que la distribución binomial con $p = 0.4$ no proporciona una buena descripción de nuestras frecuencias observadas.

SUGERENCIAS Y SUPOSICIONES

Muchas personas saben que la prueba de ji-cuadrada se puede usar como prueba de bondad de ajuste y la mayoría puede hacer los cálculos. Sin embargo, sólo unos cuantos pueden explicar la lógica de emplear la prueba con este fin, en términos de sentido común. Sugerencia: si se tiene una distribución que *se cree* que es normal, pero no se tiene la seguridad, se usa una distribución normal conocida para generar los *valores esperados* y después, con los métodos de ji-cuadrada, se ve cuánta diferencia hay entre estos valores esperados y los valores observados en una muestra tomada de la distribución que se piensa es normal. Si la diferencia es demasiado grande, la distribución no es normal.

Ejercicios 11.3

Ejercicios de autoevaluación

EA 11-3 Para el nivel de significancia de 0.10, ¿se puede concluir que las siguientes 400 observaciones siguen una distribución de Poisson con $\lambda = 3$?

Número de llegadas por hora	0	1	2	3	4	5 o más
Número de horas	20	57	98	85	78	62

EA 11-4 Después de años de trabajar en una estación de pesado para camiones, Jeff Simpson siente que el peso por camión (en miles de libras) sigue una distribución normal con $\mu = 71$ y $\sigma = 15$. Con el objeto de probar esta suposición, Jeff recolecta los siguientes datos un lunes y registra el peso de cada camión que llega a su báscula.

85	57	60	81	89	63	52	65	77	64
89	86	90	60	57	61	95	78	66	92
50	56	95	60	82	55	61	81	61	53
63	75	50	98	63	77	50	62	79	69
76	66	97	67	54	93	70	80	67	73

Si Jeff usa la prueba de bondad de ajuste de ji-cuadrada para estos datos, ¿qué concluye acerca de la distribución del peso de los camiones? (Use un nivel de significancia de 0.10 y asegúrese de establecer la hipótesis de interés.) (*Sugerencia:* use cinco intervalos igualmente probables.)

Conceptos básicos

11-14 Los datos corresponden a una distribución de frecuencias observadas. Use una distribución normal con $\mu = 5$ y $\alpha = 1.5$,

a) encuentre la probabilidad de falla en cada clase.
b) a partir del a), calcule la frecuencia esperada para cada categoría.
c) calcule el estadístico ji-cuadrada.
d) para un nivel de significancia de 0.10, ¿parece que esta distribución de frecuencias está bien descrita por la distribución normal sugerida?

Valor observado de la variable	<2.6	2.6-3.79	3.8-4.99	5-6.19	6.2-7.39	≥ 7.4
Frecuencia observada	6	30	41	52	12	9

11-15 Para un nivel de significancia de 0.05, ¿se puede concluir que los siguientes datos siguen una distribución de Poisson con $\alpha = 5$?

Número de llamadas por minuto	0	1	2	3	4	5	6	7 o más
Frecuencia de ocurrencias	4	15	42	60	89	94	52	80

Aplicaciones

11-16 Louis Armstrong, vendedor de Dillard Paper Company, debe visitar cinco cuentas diariamente. Se sugiere que la variable "ventas del señor Armstrong" puede describirse mediante la distribución binomial y con una probabilidad de venta para cada cuenta de 0.4. Dada la siguiente distribución de frecuencias del número de ventas por día del señor Armstrong, ¿podemos concluir que los datos de hecho siguen la distribución sugerida? Haga los cálculos para un nivel de significancia de 0.05.

Ventas por día	0	1	2	3	4	5
Frecuencia del número de ventas	10	41	60	20	6	3

11-17 El coordinador de computación en la escuela de administración cree que el tiempo que un estudiante de posgrado dedica a leer y escribir correos electrónicos cada día de la semana tiene una distribución normal con $\mu = 14$ y $\sigma = 5$. Para examinar esta opinión, el coordinador recolecta datos un miércoles y registra el tiempo en minutos que cada estudiante de posgrado pasa con su correo electrónico. Use la prueba de bondad de ajuste de ji-cuadrada con estos datos, ¿qué concluye acerca de la distribución del tiempo dedicado al correo electrónico? (Utilice 0.05 para el nivel de significancia y establezca con claridad sus hipótesis.) (*Sugerencia:* use cinco intervalos igualmente probables.)

8.2	7.4	9.6	12.8	22.4	6.2	8.7	9.7	12.4	10.6
1.2	18.6	3.3	15.7	18.4	12.4	15.9	19.4	12.8	20.4
12.3	11.3	10.9	18.4	14.3	16.2	6.7	13.9	18.3	19.2
14.3	14.9	16.7	11.3	18.4	18.8	20.4	12.4	18.1	20.1

11-18 Para determinar cuánto efectivo debe mantener en la bóveda, un banco quiere determinar si el depósito promedio de un cliente tiene distribución normal. Un nuevo empleado, en busca de un aumento, recolectó los siguientes datos:

Depósito	\$0-\$999	\$1,000-\$1,999	\$2,000 o más
Frecuencia observada	20	65	25

a) Calcule las frecuencias esperadas si los datos siguen una distribución normal con media de \$1,500 y desviación estándar de \$600.
b) Calcule el estadístico ji-cuadrada.
c) Establezca en forma explícita las hipótesis nula y alternativa.
d) Pruebe sus hipótesis para 0.01 y establezca su conclusión explícita.

11-19 La oficina de correos está interesada en modelar el problema de las cartas mutiladas. Se ha sugerido que cualquier carta enviada a cierta área tiene una posibilidad de 0.15 de llegar rota o mutilada. Debido a que la oficina de correos es muy grande, se puede suponer que las posibilidades de que dos cartas de sean destruidas son independientes. Se seleccionó una muestra de 310 personas y les enviaron dos cartas de prueba

a cada una. El número de personas que recibieron 0, 1 o 2 cartas rotas fue 260, 40 y 10, respectivamente. Al nivel de significancia de 0.10, ¿es razonable concluir que el número de cartas mutiladas que recibieron las personas sigue una distribución binomial con $p = 0.15$?

11-20 Una comisión de lotería estatal afirma que para un nuevo juego de lotería hay una posibilidad del 10% de obtener un premio de $1.00, una posibilidad del 5% de obtener un premio de $100.00 y una posibilidad del 85% de no obtener premio. Para probar si esta afirmación es correcta, un ganador del último juego compró 1,000 boletos para la nueva lotería. Obtuvo 87 premios de un dólar, 48 premios de 100 dólares y 865 boletos sin premio. Al nivel de significancia de 0.05, ¿es razonable la afirmación de la comisión?

11-21 Dennis Barry, administrador de un hospital, ha examinado los registros de 210 turnos de ocho horas escogidos al azar para determinar la frecuencia con la que el hospital trata casos de fractura. El número de días en que se trataron 0, 1, 2, 3, 4 y 5 o más pacientes con huesos rotos fueron 25, 55, 65, 35, 20 y 10, respectivamente. Al nivel de significancia de 0.05, ¿es razonable creer que la incidencia de casos de huesos rotos sigue una distribución de Poisson con $\lambda = 2$?

11-22 El departamento de bomberos de una ciudad grande calcula que para cualquier zona dada, durante cualquier turno de 8 horas, existe una posibilidad del 30% de recibir por lo menos un aviso de incendio. Presentamos una muestra aleatoria de avisos recibidos durante 60 días:

Núm. de turnos que recibieron avisos	0	1	2	3
Número de días	16	27	11	6

Al nivel de significancia de 0.05, ¿siguen los avisos una distribución binomial? (*Sugerencia:* combine los dos últimos grupos de modo que todas las frecuencias esperadas sean mayores de 5.)

11-23 Una diligente estudiante de estadística desea ver si es razonable suponer que unos datos de ventas se tomaron de una población normal antes de llevar a cabo una prueba de hipótesis sobre la media de las ventas. Reúne algunos datos de ventas, calcula $\bar{x} = 78$ y $s = 9$, y los tabula como sigue:

Nivel de ventas	≤65	66-70	71-75	76-80	81-85	≥86
Número de observaciones	10	20	40	50	40	40

a) ¿Es importante para la estudiante verificar si los datos tienen distribución normal? Explique su respuesta.
b) Establezca las hipótesis nula y alternativa explícitas para verificar si los datos tienen distribución normal.
c) ¿Cuál es la probabilidad (utilizando una distribución normal con $\mu = 78$ y $\sigma = 9$) de que las ventas sean menores o iguales que 65.5; estén entre 65.5 y 70.5; entre 70.5 y 75.5; entre 75.5 y 80.5; entre 80.5 y 85.5; sean mayores o iguales que 85.5?
d) Para el nivel de significancia de 0.05, ¿la distribución de frecuencias observada sigue una distribución normal?

11-24 El gerente de un supermercado lleva un registro de la llegada de clientes a las cajas para determinar cuántas debe mantener abiertas para manejar el flujo. En una muestra de 500 periodos de cinco minutos, hubo 22, 74, 115, 95, 94, 80 y 20 periodos con 0, 1, 2, 3, 4, 5 y 6 o más clientes, respectivamente. ¿Son estos datos consistentes con una distribución de Poisson con $\lambda = 3$, para un nivel de significancia de 0.05?

11-25 Un jugador profesional de béisbol, Lon Dakestraw, estuvo al bate cinco veces en cada uno de 100 juegos. Lon asegura que tiene una probabilidad de 0.4 de pegar un hit cada vez que batea. Pruebe esta afirmación al nivel de significancia de 0.05, verificando si los datos tienen una distribucióln binomial ($p = 0.4$). (*Nota:* combine clases si el número esperado de observaciones es menor que 5.)

Número de hits por juego	Número de juegos con ese número de hits
0	12
1	38
2	27
3	17
4	5
5	1

Soluciones a los ejercicios de autoevaluación

EA 11-3 H_0: Poisson con $\lambda = 3$

H_1: otra distribución

Prueba para $\alpha = 0.10$, con $6 - 1 = 5$ grados de libertad.

Llegadas/hora	0	1	2	3	4	5+
Prob. de Poisson	0.0498	0.1494	0.2240	0.2240	0.1680	0.1848
Observado	20	57	98	85	78	62
Esperado	19.92	59.76	89.60	89.60	62.20	73.92
$\frac{(f_o - f_e)^2}{f_e}$	0.000	0.127	0.788	0.236	1.736	1.922

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e} = 4.809$$

Con 5 grados de libertad y $\alpha = 0.10$, el valor crítico de χ^2 es 9.236, así que no se rechaza H_0 porque $4.809 < 9.236$. La distribución de Poisson con $\lambda = 3$ describe bien los datos.

EA 11-4 Cinco intervalos igualmente probables; probabilidad de 0.2 para cada intervalo, $50 \times 0.2 = 10$ camiones esperados por intervalo.

z	$-\infty$	-0.84	-0.25	0.25	0.84	$+\infty$
$x = 71 + 15z$	$-\infty$	58.40	67.25	74.75	83.60	$+\infty$
Observado	10	16	3	10	11	
Esperado	10	10	10	10	10	
$\frac{(f_o - f_e)^2}{f_e}$	0.0	3.6	4.9	0.0	0.1	

$$\chi^2 = 8.6$$

H_0: Los pesos de los camiones tiene distribución normal con $\mu = 71$ y $\sigma = 15$

H_1: Los pesos de los camiones tiene una distribución diferente (ya sea normal con otra μ y/o μ, o una distribución no normal)

Con $5 - 1 = 4$ grados de libertad y $\alpha = 0.10$, el valor crítico de χ^2 es 7.779, de manera que se rechaza H_0 porque $8.6 > 7.779$. Los datos no se describen bien mediante una distribución normal con $\mu = 71$ y $\mu = 15$. Jeff está equivocado.

11.4 Análisis de varianza

Función del análisis de varianza

Antes, en este capítulo, utilizamos la prueba ji-cuadrada para examinar las diferencias entre más de dos proporciones muestrales y para hacer inferencias acerca de si las muestras se tomaron de poblaciones que contenían la misma proporción. En esta sección, aprenderemos una técnica conocida como ***análisis de varianza*** **(a menudo abreviada ANOVA: *analysis of variance*), que permite probar la significancia de las diferencias entre más de dos medias muestrales.** Usando el análisis de varianza, podremos hacer inferencias acerca de si nuestras muestras se tomaron de poblaciones que tienen la misma media.

Situaciones en las que podemos utilizar ANOVA

El análisis de varianza será útil en situaciones tales como la comparación del kilometraje logrado por cinco clases diferentes de gasolina; la prueba de cuál de cuatro métodos de capacitación produce el aprendizaje más rápido; o en la comparación de los ingresos del primer año de los graduados de una media docena de escuelas de administración. En cada caso, se pueden comparar las medias de *más* de dos muestras.

Planteamiento del problema

En el problema del director de capacitación con que iniciamos el capítulo, se querían evaluar tres métodos de capacitación para determinar si había alguna diferencia en su efectividad.

Cálculo de la gran media

Después de terminar el periodo de capacitación, los especialistas en estadística de la compañía tomaron 16 nuevos empleados asignados aleatoriamente a los tres métodos de capacitación.* Con-

*Aunque en la práctica, 16 personas no constituyen una muestra estadística, hemos limitado el número para poder ilustrar las técnicas básicas del análisis de varianza y evitar cálculos tediosos.

tando los resultados de la producción de estos 16 empleados, el personal de estadística resumió los datos y calculó su producción media (vea la tabla 11-12). Ahora bien, si deseamos determinar la *gran media* o $\bar{\bar{x}}$ (la media del grupo completo de 16 empleados nuevos), podemos utilizar uno de dos métodos:

1. $$\bar{\bar{x}} = \frac{5 + 18 + 19 + 22 + 11 + 22 + 27 + 18 + 21 + 17 + 18 + 24 + 19 + 16 + 22 + 15}{16}$$

$$= \frac{304}{16}$$

$$= 19 \leftarrow \text{Gran media utilizando todos los datos}$$

2. $$\bar{\bar{x}} = (5/16)(17) + (5/16)(21) + (6/16)(19)$$

$$= \frac{304}{6}$$

$$= 19 \leftarrow \text{Gran media como promedio ponderado de las medias muestrales, utilizando los tamaños relativos de las muestras como pesos}$$

Planteamiento de las hipótesis

En este caso, la razón para utilizar análisis de varianza es decidir si estas tres muestras (una *muestra* es el pequeño grupo de empleados capacitados por cualquier método) se tomaron de poblaciones (una *población* es el número total de empleados que pudieron ser capacitados por ese método) que tienen las mismas medias. Debido a que estamos probando la efectividad de los tres métodos de capacitación, debemos determinar si las tres muestras, representadas por las medias muestrales, $\bar{x}_1 = 17$, $\bar{x}_2 = 21$ y $\bar{x}_3 = 19$, pudieron haberse tomado de poblaciones con la misma media, μ. Un planteamiento formal de las hipótesis nula y alternativa que deseamos probar sería:

Planteamiento del problema en símbolos

H_0: $\mu_1 = \mu_2 = \mu_3 \leftarrow$ Hipótesis nula

H_1: μ_1, μ_2 y μ_3 *no* son todas iguales $\leftarrow$ Hipótesis alternativa

Interpretación de los resultados

Si podemos concluir, a partir de nuestra prueba, que las medias de las muestras no difieren significativamente, podemos inferir que la selección del método de capacitación no influye en la productividad del empleado. Por otro lado, si encontramos entre las medias muestrales diferencias demasiado grandes para atribuirlas al error aleatorio de muestreo, podemos inferir que el método utilizado para capacitar a los trabajadores sí influye en su productividad. En ese caso, ajustaríamos nuestro programa de capacitación de acuerdo con los resultados.

Análisis de varianza: conceptos básicos

Suposiciones hechas en el análisis de varianza

Con el fin de utilizar el análisis de varianza, debemos suponer que cada una de las muestras se toma de una población normal y que cada una de estas poblaciones tiene la misma varianza, σ^2. Sin em-

Tabla 11-12 Producción diaria de 16 empleados nuevos

Método 1	Método 2	Método 3
		18
15	22	24
18	27	19
19	18	16
22	21	22
11	17	15
85	**105**	**114**
÷5	÷5	÷6
$17 = x_1$	$21 = x_2$	$19 = x_3$ ← Medias muestrales
$n_1 = 5$	$n_2 = 5$	$n_3 = 6$ ← Tamaños de muestra

bargo, si los tamaños de muestra son lo suficientemente grandes, no necesitamos la suposición de normalidad.

En el problema de los métodos de capacitación, la hipótesis nula establece que las tres poblaciones tienen la misma media. Si esta hipótesis es verdadera, no es necesario clasificar los datos en tres columnas, como en la tabla 11-12, y el conjunto entero de 16 mediciones de productividad puede considerarse como una muestra de una sola población. Esta población total tiene también una varianza σ^2.

El análisis de varianza está basado en una comparación de dos estimaciones diferentes de la varianza, σ^2, de nuestra población total. En este caso, podemos calcular una de esas estimaciones examinando **la varianza entre las tres medias muestrales**, que son 17, 21 y 19. La otra estimación de la varianza de la población está determinada por la **variación dentro de las tres muestras mismas**, esto es (15, 18, 19, 22, 1l), (22, 27, 18, 21, 17) y (18, 24, 19, 16, 22, 15). Entonces comparamos estas dos estimaciones de la varianza de la población. Como ambas son estimaciones de σ^2, deben tener un valor aproximadamente igual *cuando la hipótesis nula sea verdadera*. Si la hipótesis nula *no* es verdadera, estas dos estimaciones diferirán de manera considerable. Entonces, los tres pasos del análisis de varianza son:

Pasos del análisis de varianza

1. Determinar una estimación de la varianza de la población a partir de la varianza *entre las medias de las muestras*.
2. Determinar una segunda estimación de la varianza de la población a partir de la varianza *dentro de las muestras*.
2. Comparar estas dos estimaciones. Si su valor es aproximadamente igual, se acepta la hipótesis nula.

En lo que resta de esta sección, aprenderemos cómo calcular estas dos estimaciones de la varianza de la población, cómo compararlas y cómo efectuar una prueba de hipótesis e interpretar los resultados. A medida que aprendamos a hacer estos cálculos, no pierda de vista que todos ellos están basados en los tres pasos anteriores.

Cálculo de la varianza entre las medias muestrales

Búsqueda de la primera estimación de la varianza de la población

El paso 1 en el análisis de varianza indica que debemos obtener una estimación de la varianza de la población a partir de la varianza entre las tres medias de las muestras. En lenguaje estadístico, esta estimación se conoce como *varianza entre columnas*.

En el capítulo 3 utilizamos la ecuación 3-17 para calcular la varianza de la muestra:

$$\text{Varianza de la muestra} \rightarrow s^2 = \frac{\Sigma(x - \bar{x})^2}{n - 1} \qquad [3\text{-}17]$$

Primero encuentre la varianza entre las medias muestrales

Ahora bien, como estamos trabajando con tres medias muestrales y una gran media, sustituyamos $\bar{x}$ por x, $\bar{\bar{x}}$ por $\bar{x}$, y k (el número de muestras) por n, para obtener una fórmula para la varianza entre las medias de las muestras:

Varianza entre medias muestrales

$$s_{\bar{x}}^2 = \frac{\Sigma(\bar{x} - \bar{\bar{x}})^2}{k - 1} \qquad [11\text{-}4]$$

Después, encuentre la varianza de la población utilizando la varianza entre las medias muestrales

Ahora podemos regresar un momento al capítulo 6, en donde definimos el error estándar de la media como la desviación estándar de todas las muestras posibles de un tamaño dado. La fórmula para derivar el error estándar de la media es la ecuación 6-1:

Error estándar de la media (desviación estándar de todas las medias muestrales posibles de un tamaño de muestra dado) → $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ ← Raíz cuadrada del tamaño de la muestra; Desviación estándar de la población → σ [6-1]

Podemos simplificar esta ecuación con la multiplicación cruzada y luego elevando ambos lados al cuadrado para convertir la desviación estándar de la población, σ, en la varianza de la población, σ^2:

Varianza de la población

$$\sigma^2 = \sigma_{\bar{x}}^2 \times n \qquad [11\text{-}5]$$

Error estándar elevado al cuadrado (ésta es la varianza entre las medias muestrales)

Para el problema de los métodos de capacitación, no tenemos toda la información necesaria para utilizar esta ecuación y encontrar σ^2. Específicamente, no conocemos $\sigma_{\bar{x}}^2$. Sin embargo, podríamos calcular la varianza entre las medias de las muestras, $s_{\bar{x}}^2$, con la ecuación 11-4. Así, ¿por qué no sustituir $s_{\bar{x}}^2$ en lugar de $\sigma_{\bar{x}}^2$ en la ecuación 11-5 y calculamos una estimación de la varianza de la población? Esto nos da:

$$\hat{\sigma}^2 = s_{\bar{x}}^2 \times n = \frac{\Sigma n(\bar{x} - \bar{\bar{x}})^2}{k - 1}$$

Qué tamaño de muestra utilizar

Existe una pequeña dificultad al utilizar esta ecuación tal como está. En la ecuación 6-1, n representa el tamaño de la muestra, pero, ¿qué tamaño de muestra debemos usar cuando las diferentes muestras tienen diferentes tamaños? Resolvemos este problema con la ecuación 11-6, en la que cada $(\bar{x}_j - \bar{\bar{x}})^2$ se multiplica por su propia n_j.

Estimación de la varianza entre columnas

Primera estimación de la varianza de la población

$$\hat{\sigma}_b^2 = \frac{\Sigma n_j(\bar{x}_j - \bar{\bar{x}})^2}{k - 1} \qquad [11\text{-}6]$$

donde,

- $\hat{\sigma}_b^2$ = nuestra primera estimación de la varianza de la población, basada en la varianza entre las medias de las muestras (la *varianza entre columnas*)
- n_j = tamaño de la j-ésima muestra
- $\bar{x}_j$ = media muestral de la j-ésima muestra
- $\bar{\bar{x}}$ = gran media
- k = número de muestras

Ahora podemos utilizar la ecuación 11-6 y los datos de la tabla 11-12 para calcular la varianza entre columnas. En la tabla 11-13 se muestra cómo hacer dichos cálculos.

Tabla 11-13 Cálculo de la varianza entre columnas

n	x	$\bar{\bar{x}}$	$x - \bar{\bar{x}}$	$(x - \bar{\bar{x}})^2$	$n(x - \bar{\bar{x}})^2$
5	17	19	$17 - 19 = -2$	$(-2)^2 = 4$	$5 \times 4 = 20$
5	21	19	$21 - 19 = 2$	$(2)^2 = 4$	$5 \times 4 = 20$
6	19	19	$19 - 19 = 0$	$(0)^2 = 0$	$6 \times 0 = 0$
					$\Sigma n_j(x_j - \bar{\bar{x}})^2 = 40$

$$\hat{\sigma}_b^2 = \frac{\Sigma n_j(\bar{x}_j - \bar{\bar{x}})^2}{k - 1} = \frac{40}{3 - 1} \qquad [11\text{-}6]$$

$$= \frac{40}{2}$$

$$= 20 \text{ Varianza entre columnas}$$

Cálculo de la varianza dentro de las muestras

Búsqueda de la segunda estimación de la varianza de la población

El paso 2 en ANOVA requiere una segunda estimación de la varianza de la población, basada en la varianza dentro de las muestras. En términos estadísticos, se le puede llamar *varianza dentro de columnas*. El problema de capacitación tiene tres muestras de cinco o seis elementos cada una. Podemos calcular la varianza dentro de estas tres muestras usando la ecuación 3-17:

$$\text{Varianza de la muestra} \longrightarrow s^2 = \frac{\Sigma n(x - \bar{x})^2}{n - 1} \qquad [3\text{-}17]$$

Dado que hemos supuesto que la varianza de nuestras tres poblaciones es la misma, podemos utilizar cualquiera de las tres varianzas muestrales (s_1^2, s_2^2 o s_3^2) como la segunda estimación de la varianza de la población. En términos estadísticos, podemos obtener una mejor estimación de la varianza de la población mediante un promedio ponderado de las tres varianzas de muestra. La fórmula general para esta segunda estimación de σ^2 es

Estimación de la varianza dentro de columnas

$$\text{Segunda estimación de la varianza de la población} \longrightarrow \hat{\sigma}_w^2 = \Sigma\left(\frac{n_j - 1}{n_T - k}\right)s_j^2 \qquad [11\text{-}7]$$

donde,

- $\hat{\sigma}_w^2$ = nuestra segunda estimación de la varianza de la población, basada en las varianzas dentro de las muestras (la *varianza dentro de columnas*)
- n_j = tamaño de la *j*-ésima muestra
- s_j^2 = varianza muestral de la *j*-ésima muestra
- k = número de muestras
- $n_T = \Sigma n_j$ = tamaño de la muestra total

Uso de toda la información disponible

Esta fórmula utiliza toda la información que tenemos a nuestra disposición, no nada más una parte de ella. De haber tenido siete muestras en lugar de tres, habríamos tomado un promedio ponderado de las siete. Más adelante explicaremos los pesos usados en la ecuación 11-7. En la tabla 11-14 se ilustra cómo calcular esta segunda estimación de la varianza de la población, utilizando las varianzas dentro de las tres muestras.

Prueba de hipótesis *F*: cálculo e interpretación del estadístico *F*

Búsqueda del cociente *F*

En el paso 3 de ANOVA se comparan estas dos estimaciones de la varianza de la población mediante el cálculo de su cociente como sigue:

$$F = \frac{\begin{array}{c}\text{primera estimación de la varianza de la} \\ \text{población basada en la varianza entre las medias muestrales}\end{array}}{\begin{array}{c}\text{segunda estimación de la varianza de la población} \\ \text{basada en las varianzas dentro de las muestras}\end{array}} \qquad [11\text{-}8]$$

Si sustituimos con la terminología estadística, en el numerador y el denominador de este cociente, la ecuación 11-8 se convierte en:

Estadístico *F*

$$F = \frac{\text{varianza entre columnas}}{\text{varianza dentro de columnas}} = \frac{\hat{\sigma}_b^2}{\hat{\sigma}_w^2} \qquad [11\text{-}9]$$

Tabla 11-14

Cálculo de las varianzas dentro de las muestras y la varianza dentro de columnas

Método de capacitación 1 Media muestral: $x = 17$		Método de capacitación 2 Media muestral: $x = 21$		Método de capacitación 3 Media muestral: $x = 19$	
$x - x$	$(x - x)^2$	$x - x$	$(x - x)^2$	$x - x$	$(x - x)^2$
$15 - 17 = -2$	$(-2)^2 = 4$	$22 - 21 = 1$	$(1)^2 = 1$	$18 - 19 = -1$	$(-1)^2 = 1$
$18 - 17 = 1$	$(1)^2 = 1$	$27 - 21 = 6$	$(6)^2 = 36$	$24 - 19 = 5$	$(5)^2 = 25$
$19 - 17 = 2$	$(2)^2 = 4$	$18 - 21 = -3$	$(-3)^2 = 9$	$19 - 19 = 0$	$(0)^2 = 0$
$22 - 17 = 5$	$(5)^2 = 25$	$21 - 21 = 0$	$(0)^2 = 0$	$16 - 19 = -3$	$(-3)^2 = 9$
$11 - 17 = -6$	$(-6)^2 = 36$	$17 - 21 = -4$	$(-4)^2 = 16$	$22 - 19 = 3$	$(3)^2 = 9$
	$\Sigma(x - x)^2 = 70$		$\Sigma(x - x)^2 = 62$	$15 - 19 = -4$	$(-4)^2 = 16$
					$\Sigma(x - x)^2 = 60$
$\frac{\Sigma(x - \bar{x})^2}{n - 1} = \frac{70}{5 - 1} = \frac{70}{4}$		$\frac{\Sigma(x - \bar{x})^2}{n - 1} = \frac{62}{5 - 1} = \frac{62}{4}$		$\frac{\Sigma(x - \bar{x})^2}{n - 1} = \frac{70}{6 - 1} = \frac{60}{5}$	
Varianza de la muestra $\rightarrow s_1^2 = 17.5$		Varianza de la muestra $\rightarrow s_2^2 = 15.5$		Varianza de la muestra $\rightarrow s_3^2 = 12.0$	

Y:
$$\hat{\sigma}^2 = \sum\left(\frac{n_j - 1}{n_T - k}\right)s_j^2 = (4/13)(17.5) + (4/13)(15.5) + (5/13)(12.0) \qquad [11\text{-}7]$$
$$= \frac{192}{13}$$
$$= 14.769 \leftarrow$$ Segunda estimación de la varianza de la población basado en las varianzas dentro de las muestras (la varianza dentro de columnas)

Ahora podemos encontrar el cociente *F* para el problema del método de capacitación que hemos estado manejando:

$$F = \frac{\text{varianza entre columnas}}{\text{varianza dentro de columnas}} = \frac{\hat{\sigma}_b^2}{\hat{\sigma}_w^2} \qquad [11\text{-}9]$$
$$= \frac{20}{14.769}$$
$$= 1.354 \leftarrow \text{Cociente } F$$

Interpretación del cociente *F*

Una vez encontrado el cociente *F*, 1.354, ¿cuál es su interpretación? Primero examinamos el denominador, que está basado en la varianza dentro de las muestras. El denominador es un buen estimador de σ^2 (la varianza de la población) ya sea que la hipótesis nula sea verdadera o no. ¿Qué sucede con el numerador? Si la hipótesis nula de que los métodos de capacitación tienen el mismo efecto en la producción es verdadera, entonces el numerador, o la variación entre las medias de la muestras de los tres métodos, es también una buena estimación de σ^2 (la varianza de la población). Como resultado, **el denominador y el numerador deben ser aproximadamente iguales si la hipótesis nula es verdadera.** Cuanto más cercano a 1 esté el cociente *F*, más nos inclinamos a aceptar la hipótesis nula. Al contrario, conforme el cociente *F* crece, nos inclinaremos más a rechazar la hipótesis nula y a aceptar la alternativa (de que existe una diferencia en los efectos sobre la producción de los tres métodos de capacitación).

Dentro de poco aprenderemos una manera más formal de decidir cuándo aceptar o rechazar la hipótesis nula. Pero incluso en este momento, debe entender la lógica básica que apoya el *estadístico F*. **Cuando las poblaciones no son las mismas, la varianza entre columnas (derivada a partir de la varianza entre las medias muestrales) tenderá a ser mayor que la varianza dentro de columnas (derivada a partir de la varianza dentro de las muestras), y el valor de *F* tenderá a ser grande. Esto nos conducirá a rechazar la hipótesis nula.**

La distribución *F*

Descripción de la distribución *F*

Como otros estadísticos que hemos estudiado, si la hipótesis nula es verdadera, entonces el estadístico *F* tiene una distribución de muestreo específica. Al igual que las distribuciones *t* y ji-cuadrada, la distribución *F* es en realidad una familia completa de distribuciones, tres de las cuales se observan en la figura 11-8. Note que cada una está identificada por un *par* de grados de libertad, a diferencia de las distribuciones *t* y ji-cuadrada, que solamente tienen un valor para el número de grados de libertad. **La primera cantidad se refiere a los grados de libertad del numerador del cociente *F*; la segunda, a los grados de libertad del denominador**.

Como podemos ver en la figura 11-8, la distribución *F* tiene una sola moda. La forma específica de una distribución *F* depende del número de grados de libertad tanto del numerador como del denominador del cociente *F*. Pero, en general, la distribución está sesgada a la derecha y tiende a hacese más simétrica conforme aumenta el número de grados de libertad en el numerador y el denominador.

Uso de la distribución *F*: grados de libertad

Cálculo de los grados de libertad

Como se mencionó, cada distribución *F* tiene una par de grados de libertad, uno para el numerador del cociente *F* y el otro para el denominador. ¿Cómo podemos calcularlos?

Búsqueda de los grados de libertad del numerador

Primero, piense en el numerador, la varianza entre columnas. En la tabla 11-13 utilizamos tres valores de $\bar{x} - \bar{\bar{x}}$, uno para cada muestra, para calcular $\Sigma n_j(\bar{x}_j - \bar{\bar{x}})^2$. Una vez conocidos dos de estos valores de $\bar{x} - \bar{\bar{x}}$, el tercero queda *automáticamente determinado* y no se puede determinar libremente. Así, se pierde un grado de libertad cuando calculamos la varianza entre columnas, y el número de grados de libertad para el numerador del cociente *F* siempre es una unidad menor que el número de muestras. La regla, entonces, es:

Grados de libertad del numerador

$$\text{Número de grados de libertad en el } \textit{numerador} \text{ del cociente } F = (\text{número de muestras} - 1) \qquad [11\text{-}10]$$

Búsqueda de los grados de libertad del denominador

Ahora, ¿qué pasa con el denominador? Observe un momento la tabla 11-14. Calculamos las varianzas dentro de las muestras, y utilizamos las tres muestras. Para la *j*-ésima muestra, usamos n_j valores de $(x - \bar{x}_j)$ para calcular $\Sigma(x - \bar{x}_j)^2$ para esa muestra. Cuando tenemos todos los valores de $(x - \bar{x}_j)$, excepto uno, el último queda *determinado automáticamente* y no se podría especificar libremente. En consecuencia, perdimos un grado de libertad en los cálculos de *cada* muestra, lo que deja 4, 4 y 5 grados de libertad en las muestras. Como tenemos tres muestras, nos quedamos con

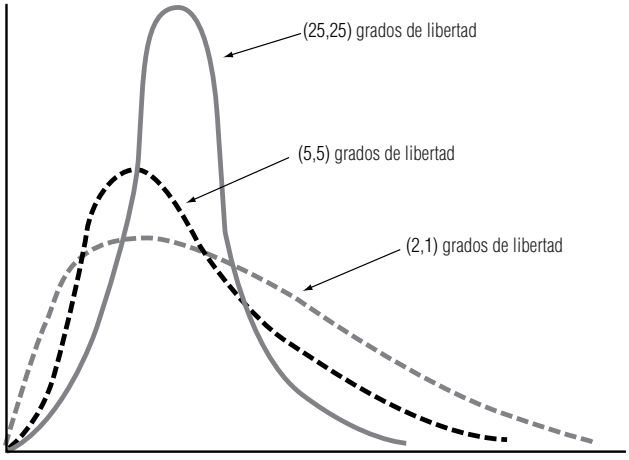


FIGURA 11-8

Tres distribuciones *F* (el primer valor entre paréntesis es igual al número de grados de libertad del numerador del cociente *F*; el segundo, al número de grados de libertad del denominador)

$4 + 4 + 5 = 13$ grados de libertad (que también pueden calcularse como $5 + 5 + 6 - 3 = 13$). Podemos establecer la regla de la siguiente manera:

Grados de libertad del denominador

$$\text{Número de grados de libertad en el } \textit{denominador} \text{ del cociente } F = \Sigma(n_j - 1) = n_T - k \qquad [11\text{-}11]$$

donde,

- n_j = tamaño de la j-ésima muestra
- k = número de muestras
- $n_T = \Sigma n_j$ = tamaño de la muestra total

Ahora vemos que el peso asignado a s_j^2 en la ecuación 11-7 es justo la fracción del número total de grados de libertad del denominador del cociente F.

Uso de la tabla *F*

Para llevar a cabo pruebas de hipótesis F debemos utilizar una tabla F, en la cual las columnas representan el número de grados de libertad del numerador y los renglones el número de grados de libertad del denominador. Existen tablas separadas para cada nivel de significancia.

Suponga que estamos probando una hipótesis al nivel de significancia de 0.01, usando la distribución F. Nuestros grados de libertad son 8 para el numerador y 11 para el denominador. En este caso, vamos a la tabla 6(b) del apéndice. En el cuerpo de dicha tabla, el valor apropiado para 8 y 11 grados de libertad es 4.74. Si nuestro valor calculado de F excede este valor de la tabla, rechazamos la hipótesis nula. Si no es mayor, la aceptamos.

Prueba de hipótesis

Búsqueda de el estadístico *F* y los grados de libertad

Ahora podemos probar nuestra hipótesis de que los tres métodos de capacitación producen resultados idénticos, utilizando el material desarrollado hasta este punto. Empecemos por revisar cómo calcular el cociente F:

$$F = \frac{\text{primera estimación de la varianza de población basada en la varianza entre las medias muestrales}}{\text{segunda estimación de la varianza de la población basada en las varianzas dentro de las muestras}} \qquad [11\text{-}8]$$

$$= \frac{20}{14.769}$$

$$= 1.354 \leftarrow \text{Estadístico } F$$

A continuación, calculamos el número de grados de libertad del numerador del cociente F, con la ecuación 11-10 como sigue:

$$\text{Número de grados de libertad en el } \textit{numerador} \text{ del cociente } F = (\text{número de muestras} - 1) \qquad [11\text{-}10]$$

$$= 3 - 1$$

$$= 2 \leftarrow \text{Grados de libertad en el numerador}$$

Y podemos calcular los grados de libertad del denominador del cociente F, utilizando la ecuación 11-11:

Número de grados de libertad en el *denominador* del cociente F

$$= \Sigma(n_j - 1) = n_T - k \quad [11\text{-}11]$$
$$= (5 - 1) + (5 - 1) + (6 - 1)$$
$$= 4 + 4 + 5$$
$$= 13 \leftarrow \text{Grados de libertad en el denominador}$$

Búsqueda del límite de la región de aceptación

Interpretación de los resultados

Suponga que el director de capacitación desea probar al nivel de significancia de 0.05 la hipótesis de que no existen diferencias entre los tres métodos. Podemos buscar en la tabla 6(a) del apéndice para 2 grados de libertad en el numerador y 13 en el denominador. El valor que encontramos es 3.81. La figura 11-9 ilustra esta prueba de hipótesis con una gráfica. La región sombreada es el nivel de significancia. El valor encontrado en la tabla, 3.81, establece el límite superior de la región de aceptación. Como el valor de la muestra calculado para F, 1.354, se encuentra dentro de la región de aceptación, aceptamos la hipótesis nula y concluimos que, según la información de las muestras que poseemos, no existen diferencias significativas en los efectos de los tres métodos de capacitación sobre la productividad de un empleado.

Precauciones acerca del uso de la prueba *F*

Use tamaños de muestra grandes

Como se estableció, nuestros tamaños de muestra en este problema son demasiado pequeños para poder llegar a inferencias válidas con respecto a la efectividad de los métodos de capacitación. Escogimos muestras pequeñas para explicar la lógica del análisis de varianza sin tener que efectuar cálculos tediosos. En la práctica real, nuestra metodología sería la misma, pero las muestras serían más grandes.

Controle todos los factores, menos el que se está probando

En nuestro ejemplo, hemos supuesto la ausencia de muchos factores que podrían haber afectado nuestras conclusiones. Por ejemplo, aceptamos como un hecho que todos los empleados nuevos de la muestra demostraron la misma aptitud para el aprendizaje, lo cual puede ser o no cierto. Supusimos que todos los instructores de los tres métodos tienen la misma habilidad para enseñar y manejar a las personas, lo cual puede no ser cierto. Y supusimos que el personal de estadística de la compañía reunió los datos de productividad durante periodos de trabajo similares en cuanto a la hora del día, el día de la semana, la época del año, etc. Para poder tomar decisiones significativas basadas en el análisis de varianza, necesitamos tener la certeza de que todos los factores mencionados están controlados de manera efectiva.

Una prueba para un solo factor

Por último, note que solamente estudiamos el análisis de varianza en *un sentido*, o de un factor. El problema examinó el efecto del tipo de método de capacitación sobre la productividad de los empleados y nada más. Si deseáramos medir el efecto de dos factores, como el programa de capacitación y la edad del empleado, necesitaríamos usar análisis de varianza en dos sentidos, un método estadístico que se estudia en textos más avanzados.

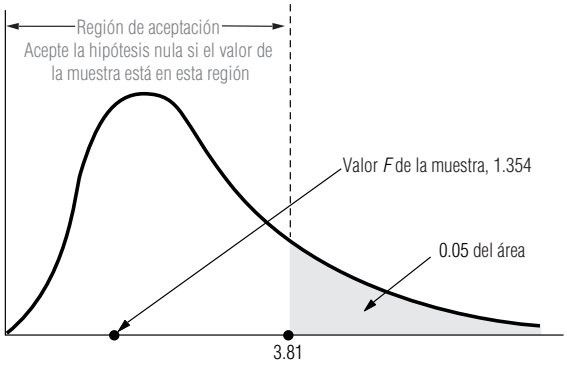


FIGURA 11-9

Prueba de hipótesis al nivel de significancia de 0.05, utilizando la distribución *F*; indica la región de aceptación y el valor *F* de la muestra

Uso de la computadora para el análisis de varianza

Uso del paquete SPSS para ANOVA

Una vez más, repetimos que usamos muestras pequeñas en el ejemplo de ANOVA, con el fin de explicar la lógica del método sin abrumarnos con cálculos tediosos. Para un problema más realista, sería muy conveniente usar las rutinas de ANOVA que se pueden encontrar en todos los paquetes estadísticos de uso común. Para comparar uno de estos paquetes con el análisis hecho a mano, la figura 11-10 presenta la salida del paquete SPSS, usado para analizar los datos del problema del método de capacitación.

Veamos la columna de la tabla ANOVA producida por SPSS que tiene el encabezado "MEDIA CUADRADA". En el renglón correspondiente a "MODELO", esta columna contiene el valor 20.000, que reconocemos como la varianza entre columnas calculada en la tabla 11-13. En el renglón de "ERROR", de la figura 11-10 encontramos el valor 14.769, que es la varianza dentro de columnas calculada en la tabla 11-14. Note también la columna con "GL" (que significa grados de libertad). Nos dice que la MEDIA CUADRADA DEL MODELO (la varianza entre columnas) tiene 2 grados de libertad, y que el ERROR MEDIO CUADRADO (la varianza dentro de columnas) tiene 13 grados de libertad.

La última línea de la salida de SPSS da el valor del estadístico F, $F = 1.35$, y el valor p, 0.2923, que es la probabilidad de obtener un estadístico F igual o mayor que 1.35 si H_0 es verdadera. Como el valor p es mayor que nuestro nivel de significancia, $\alpha = 0.05$, de nuevo concluimos que no podemos rechazar H_0. Con base en la evidencia de la muestra, estos tres métodos de capacitación no parecen tener efectos significativamente diferentes sobre la productividad de los empleados.

La figura 11-11 ilustra el uso de Minitab para realizar la misma prueba. Minitab proporciona, en esencia, la misma tabla ANOVA, junto con n, $\bar{x}$ y s para las tres muestras.

Ahora que hemos visto cómo interpretar los resultados obtenidos con SPSS y Minitab, analicemos un ejemplo mucho más realista. A pesar de la conclusión dada en la sección 9.7, de que los es-

EJEMPLO DEL USO DE SPSS PARA ANOVA
¿LA PRODUCTIVIDAD DEPENDE DEL MÉTODO DE CAPACITACIÓN?

PROCEDIMIENTO DE ANÁLISIS DE VARIANZA

VARIABLE DEPENDIENTE: UNIDADES UNIDADES PROD. POR EL EMPLEADO

FUENTE	GL	SUMA DE CUADRADOS	MEDIA CUADRADA
MODELO	2	40.00000000	20.00000000
ERROR	13	192.00000000	14.76923077
TOTAL CORREGIDO	15	232.00000000	

F DEL MODELO = 1.35 PR > F = 0.2923

FIGURA 11-10

Salida de SPSS para el problema de la capacitación de empleados

Análisis de varianza en un sentido

Análisis de varianza

Fuente	GL	SC	MC	F	p
Factor	2	40.0	20.0	1.35	0.292
Error	13	192.0	24.8		
Total	15	232.0			

Nivel	N	Media	DesvEst
MÉTODO 1	5	17.000	4.183
MÉTODO 2	5	21.000	3.937
MÉTODO 3	6	19.000	3.464

DesvEst agrupada = 3.843

FIGURA 11-11

Salida de Minitab para el problema de capacitación de empleados

tudiantes que tomaron el curso de estadística con profesores no obtenían significativamente mejores resultados en el examen final que los estudiantes que tomaron el curso con ayudantes de profesor, todavía se recibieron quejas. "Estoy en la clase del señor Jackson, y mis amigos que están en el grupo del profesor Rubín han aprendido mucho más que yo", era una queja típica. No preguntamos entonces si, tal vez, había *diferencias significativas entre las secciones de un mismo curso*, incluso si los ayudantes de profesor *como grupo* no eran significativamente diferentes de los profesores *como grupo*.

Utilizamos ANOVA de Minitab para verificar lo anterior. El planteamiento formal de nuestras hipótesis fue:

H_0: Las seis μ son iguales (no hay diferencia entre secciones)

H_1: Las seis μ no son iguales (los cursos difieren de manera significativa)

Conclusiones acerca de las quejas de los estudiantes

La figura 11-12 contiene los resultados de este análisis. El valor calculado del estadístico F es 1.75, y la probabilidad de observar un valor tan grande de F si H_0 es verdadera (el valor p para esta prueba) es 0.126. Con un valor p tan grande, debemos aceptar H_0 y llegar a la conclusión de que no hay diferencias significativas en el desempeño de los estudiantes de las seis secciones en el examen final.

Uso de ANOVA con los datos de ingresos

Usemos Minitab para hacer un análisis de varianza con los datos de ingresos de la tabla 11 del apéndice. Recuerde que en la sección 9.7 se usó Minitab para probar si los cambios de un año a otro en los ingresos del último trimestre de las acciones de la Bolsa de Valores de Nueva York (NYSE)

Análisis de varianza en un sentido

Análisis de varianza para FINAL

Fuente	GL	SC	MC	F	p
SECCIÓN	5	859.4	171.9	1.75	0.126
Error	193	18996.8	98.4		
Total	198	19856.2			

Nivel	N	Media	DesvEst
1	27	45.741	10.679
2	46	44.761	11.900
3	37	49.081	7.365
4	26	44.923	8.064
5	36	44.333	10.373
6	27	42.111	9.435

DesvEst agrupada = 9.921

FIGURA 11-12

Salida de Minitab para ANOVA de las calificaciones en exámenes finales

Análisis de varianza en un sentido

Análisis de varianza para CAMBIO

Fuente	GL	SC	MC	F	p
MERCADO	2	1.421	0.711	0.88	0.415
Error	221	177.906	0.805		
Total	223	179.327			

Nivel	N	Media	DesvEst
1	111	-0.1120	0.5195
2	38	0.0876	0.9121
3	75	0.0156	1.2598

DesvEst agrupada = 0.8972

FIGURA 11-13

Uso de Minitab para calcular la ANOVA de cambios en ganancias

tenían medias diferentes que los cambios en los ingresos del último trimestre del American Stock Exchange (ASE) y de otras bolsas. Se agruparon los últimos dos porque en ese punto todavía no sabíamos cómo comparar más de dos medias. Ahora se verán los tres grupos por separado y se usará análisis de varianza para ver si las tres medias difieren de manera significativa.

H_0: $\mu_{OTRO} = \mu_{ASE} = \mu_{NYSE}$ (sin diferencias por transacción)

H_1: Las μ no son iguales (las transacciones difieren significativamente)

Interpretación de los resultados

Se usó el comando ONEWAY (un sentido) en Minitab para realizar este análisis de varianza. Recuerde que la variable MERCADO es 1 para OTRO, 2 para ASE y 3 para NYSE. Los resultados se dan en la figura 11-13. El valor calculado del estadístico F es 0.88 y el valor p para probar la hipótesis es 0.415. Como este valor p es más grande que todos los niveles de significancia usuales ($\alpha = 0.10$, 0.05, 0.01, etc.), no se puede rechazar H_0; se concluye que los valores medios de los cambios de un año a otro en los ingresos del último trimestre para los tres mercados de valores no difieren de manera significativa uno de otro.

SUGERENCIAS Y SUPOSICIONES

El análisis de varianza se centra en probar si tres o más muestras se han obtenido de poblaciones que tienen la misma media. El análisis de varianza es importante en investigaciones como la evaluación de nuevos medicamentos, donde deben medirse, en un solo estudio, los efectos de dosis, la frecuencia de medicación, los efectos de otras drogas y las diferencias entre pacientes. Una estimación se obtiene de la varianza entre las medias de las muestras, la otra de la varianza dentro de las muestras mismas. Si son aproximadamente iguales, se tienen altas posibilidades de que las muestras vengan de la misma población. Advertencia: es vital no hacer a un lado el sentido común al interpretar los resultados. Mientras que quizá sea cierto que un estudio puede identificar diferencias en las preferencias de marca del café instantáneo que se aplican a la compra de café entre semana en la mañana, es difícil decir qué debe hacer la compañía de café con esta información.

Ejercicios 11.4

Ejercicios de autoevaluación

EA 11-5 Un estudio compara los efectos sobre las ventas de 4 promociones de un mes en el punto de venta. Presentamos las ventas unitarias de 5 tiendas que utilizaron las 4 promociones en meses distintos:

Muestras gratis	78	87	81	89	85
Regalo de un paquete	94	91	87	90	88
Descuento	73	78	69	83	76
Reembolso por correo	79	83	78	69	81

a) Calcule las ventas unitarias medias para cada promoción y luego determine la gran media.
b) Estime la varianza de la población utilizando la varianza entre columnas (ecuación 11-6).
c) Estime la varianza de la población utilizando la varianza dentro de columnas calculada a partir de la varianza dentro de muestras.
d) Calcule el cociente F. Al nivel de significancia de 0.01, ¿las promociones producen diferentes efectos sobre las ventas?

EA 11-6 Una compañía de investigación diseñó tres sistemas diferentes para limpiar manchas de aceite. La siguiente tabla contiene los resultados, medidos por el área (en metros cuadrados) que se limpia en una hora. Los datos se encontraron al probar cada método en varios ensayos. ¿Tienen la misma efectividad los tres métodos? Use un nivel de significancia de 0.05.

Sistema A	55	60	63	56	59	55
Sistema B	57	53	64	49	62	
Sistema C	66	52	61	57		

Aplicaciones

■ **11-26** Un estudio compara el número de horas de alivio que proporcionan cinco marcas de antiácidos administrados a 25 personas diferentes, cada una con acidez estomacal considerada fuerte. Los resultados son los siguientes:

Marca	A	B	C	D	E
	4.4	5.8	4.8	2.9	4.6
	4.6	5.2	5.9	2.7	4.3
	4.5	4.9	4.9	2.9	3.8
	4.1	4.7	4.6	3.9	5.2
	3.8	4.6	4.3	4.3	4.4

a) Calcule el número medio de horas de alivio para cada marca y determine la gran media.
b) Estime la varianza de la población usando la varianza entre columnas (ecuación 11-6).
c) Estime la varianza de la población usando la varianza dentro de columnas calculada a partir de la varianza dentro de las muestras.
d) Calcule el cociente *F*. Para un nivel de significancia de 0.05, ¿las marcas producen cantidades significativamente diferentes de alivio a las personas con acidez estomacal fuerte?

■ **11-27** Se compararon tres métodos de capacitación para ver si los empleados tienen una mayor productividad después de capacitarse. Los datos que se presentan a continuación son medidas de la productividad de los individuos capacitados por cada método.

Método 1	45	40	50	39	53	44
Método 2	59	43	47	51	39	49
Método 3	41	37	43	40	52	37

Al nivel de significancia de 0.05, ¿los tres métodos de entrenamiento llevan a diferentes niveles de productividad?

■ **11-28** Los datos siguientes indican el número de solicitudes de pago de seguro procesadas diariamente por un grupo de cuatro empleados de aseguradoras observados durante cierto número de días. Pruebe la hipótesis de que las solicitudes medias de los empleados por día son las mismas. Utilice un nivel de significancia de 0.05.

Empleado 1	15	17	14	12		
Empleado 2	12	10	13	17		
Empleado 3	11	14	13	15	12	
Empleado 4	13	12	12	14	10	9

■ **11-29** Dadas las siguientes mediciones de cuatro muestras, ¿podemos concluir que vienen de poblaciones que tienen el mismo valor medio? Utilice el nivel de significancia de 0.01.

Muestra 1	16	21	24	28	29	
Muestra 2	29	18	20	19	30	21
Muestra 3	14	15	21	19	28	17
Muestra 4	21	28	20	22	18	

■ **11-30** El gerente de una línea de ensamble de una planta manufacturera de relojes decidió estudiar de qué manera las diferentes velocidades de la banda transportadora afectan la tasa de unidades defectuosas producidas en un turno de 8 horas. Para ello, corrió la banda a 4 velocidades distintas en 5 turnos de 8 horas cada uno y registró el número de unidades defectuosas encontradas al final de cada turno. Los resultados del estudio son los siguientes:

	Unidades defectuosas por turno		
Velocidad 1	**Velocidad 2**	**Velocidad 3**	**Velocidad 4**
37	27	32	35
35	32	36	27
38	32	33	33
36	34	34	31
34	30	40	29

a) Calcule el número medio de unidades defectuosas, $\bar{x}$, para cada velocidad; luego determine la gran media, $\bar{\bar{x}}$.
b) Utilizando la ecuación 11-6, estime la varianza de la población (la varianza entre columnas).
c) Calcule las varianzas *dentro* de las muestras y estime la varianza de la población basándose en estas varianzas (la varianza dentro de columnas).
d) Calcule el cociente *F*. Al nivel 0.05 de significancia, ¿las cuatro velocidades de la banda transportadora producen la misma tasa media de relojes defectuosos por turno?

■ **11-31** Estamos interesados en probar la diferencia en sabor de tres salsas condimentadas: A, B y C. Para cada producto se tomó una muestra de 25 personas. Cada persona calificó el producto de -3 (terrible) a $+3$ (excelente). El paquete SPSS produjo el siguiente informe:

```
          PROCEDIMIENTO DE ANÁLISIS DE VARIANZA

VARIABLE DEPENDIENTE:   CALIFICACIÓN (-3 A +3)
FUENTE                   GL   SUMA DE CUADRADOS      MEDIA CUADRADA

MODELO                    2         15.68                  7.84
ERROR                    72          94.4             1.31111111
TOTAL CORREGIDO          74        110.08
F DEL MODELO =         5.98                          PR > F = 0.004
```

a) Establezca las hipótesis nula y alternativa explícitas.
b) Pruebe sus hipótesis con la salida de SPSS. Use $\alpha = 0.05$.
c) Establezca una conclusión explícita.

■ **11-32** El supervisor de seguridad de una tienda departamental grande desea saber si el personal de seguridad sorprende a una cantidad relativamente mayor de ladrones durante la temporada navideña que en las semanas anteriores o posteriores. Reunió datos correspondientes al número de ladrones aprehendidos en la tienda durante los meses de noviembre, diciembre y enero, durante los seis años anteriores. La información es:

	Número de ladrones					
Noviembre	43	37	59	55	38	48
Diciembre	54	41	48	35	50	49
Enero	36	28	34	41	30	32

Al nivel de significancia de 0.05, ¿es el número medio de ladrones sorprendidos el mismo durante estos tres meses?

■ **11-33** Un curso de introducción a la economía se ofrece en 3 secciones, cada una con diferente instructor. Las calificaciones finales del semestre de primavera se presentan en la tabla. ¿Existe una diferencia significativa en los promedios de calificaciones dadas por los instructores? Establezca las pruebas de hipótesis adecuadas para $\alpha = 0.01$.

Sección 1	Sección 2	Sección 3
98.4	97.6	94.5
97.6	99.2	92.3
84.7	82.6	92.4
88.5	81.2	82.3
77.6	64.5	62.6
84.3	82.3	68.6
81.6	68.4	92.7
88.4	75.6	82.3
95.1		91.2
90.4		92.6
89.4		87.4
65.6		
94.5		
99.4		
68.7		
83.4		

■ **11-34** Los fabricantes de chips de silicio requieren los llamados cuartos limpios, donde el aire se filtra de manera especial para mantener el número de partículas de polvo al mínimo. La Outel Corporation desea asegurarse de que cada uno de sus cinco cuartos limpios tenga el mismo número de partículas de polvo. Se tomaron cinco muestras de aire en cada cuarto. Se midió el "nivel de polvo" en una escala de 1 (bajo) a 10 (alto). Al nivel de significancia de 0.05, ¿tienen los cuartos el mismo nivel promedio de polvo?

	Nivel de polvo (1 a 10)				
Cuarto 1	5	6.5	4	7	6
Cuarto 2	3	6	4	4.5	3
Cuarto 3	1	1.5	3	2.5	4
Cuarto 4	8	9.5	7	6	7.5
Cuarto 5	1	2	3.5	1.5	3

■ **11-35** Una compañía maderera está preocupada por saber cómo las tasas de interés crecientes afectan a la construcción de casas nuevas en el área. Para explorar esta cuestión, la compañía ha reunido datos con respecto a nuevas construcciones durante los tres trimestres pasados en tres de los municipios circundantes. Esta información se presenta en la siguiente tabla. Al nivel de significancia de 0.05, ¿existen diferencias en el número de nuevas construcciones de casas durante los tres trimestres?

Trimestre 1	41	53	54	55	43
Trimestre 2	45	51	48	43	39
Trimestre 3	34	44	46	45	51

■ **11-36** La compañía Genes-and-Jeans, Inc., ofrece clones de cuatro marcas famosas de pantalones jeans: Generic, ADN, ARN y Oops. La tienda desea ver si existen diferencias en el número de pantalones vendidos de cada marca. El gerente ha contado los pantalones vendidos de cada marca en varios días. Al nivel de significancia de 0.05, ¿son iguales las ventas de las cuatro marcas?

	Pantalones vendidos					
Generic	17	21	13	27	12	
ADN	27	13	29	9		
ARN	13	15	17	23	10	21
Oops	18	25	15	27	12	

■ **11-37** La Oficina de Contabilidad del Gobierno (OCG) de Estados Unidos está interesada en ver si las oficinas de tamaño parecido gastan cantidades similares en personal y equipo. (Las oficinas que gastan más tendrán una auditoría especial.) Se examinaron los gastos mensuales de tres oficinas: una de ellas en el Departamento de Agricultura, otra en el Departamento de Estado y la última en el Departamento del Interior. Los datos se presentan en la tabla. Al nivel de significancia de 0.01, ¿existen diferencias en los gastos de las distintas oficinas?

	Gastos mensuales (en miles de dólares) durante algunos meses					
Agricultura	10	8	11	9	12	
Estado	15	9	8	10	13	13
Interior	8	16	12			

■ **11-38** En la ciudad de Bigville, una cadena de comida rápida está adquiriendo una mala reputación debido a que tardan mucho en servir a los clientes. Como la cadena tiene cuatro restaurantes en esa ciudad, quiere saber si los cuatro restaurantes tienen el mismo tiempo promedio de servicio. Uno de los dueños de la cadena ha decidido visitar cada local y registrar el tiempo de servicio para 5 clientes escogidos al azar. En sus cuatro visitas al medio día registró los siguientes tiempos de servicio en minutos:

Restaurante 1	3	4	5.5	3.5	4
Restaurante 2	3	3.5	4.5	4	5.5
Restaurante 3	2	3.5	5	6.5	6
Restaurante 4	3	4	5.5	2.5	3

a) Utilice un nivel de significancia de 0.05, ¿todos los restaurantes tienen el mismo tiempo medio de servicio?
b) Según sus resultados, ¿deberá el dueño hacer algunas recomendaciones a cualquiera de los administradores de los restaurantes?

Soluciones a los ejercicios de autoevaluación

EA 11-5 a)

	Gratis	Regalo	Descuento	Reembolso
	78	94	73	79
	87	91	78	83
	81	87	69	78
	89	90	83	69
	85	88	76	81
Σx	**420**	**450**	**379**	**390**
n	5	5	5	5
$\bar{x}$	84	90	75.8	78
Σx^2	35,360	40,530	28,839	30,536
s^2	20	7.5	27.7	29

$$\text{Gran media} = \bar{\bar{x}} = \frac{420 + 450 + 379 + 390}{20} = 81.95$$

b) $$\hat{\sigma}_b^2 = \frac{\Sigma n_j (\bar{x}_j - \bar{\bar{x}})^2}{k - 1} = \frac{5[(84 - 81.95)^2 + (90 - 81.95)^2 + (75.8 - 81.95)^2 + (78 - 81.95)2]}{4 - 1}$$

$$= \frac{612.15}{3} = 204.05$$

c) $$\hat{\sigma}_w^2 = \sum\left(\frac{n_j - 1}{n_T - k}\right) s_j^2 = \frac{4(20 + 7.5 + 27.7 + 29)}{20 - 4} = \frac{336.8}{16} = 21.05$$

d) $$F = \frac{204.05}{21.05} = 9.69$$

Con 3 grados de libertad en el numerador, 16 grados de libertad en el denominador y $\alpha = 0.01$, el valor crítico de F es 5.29, de manera que se rechaza H_0 porque 9.69 > 5.29. Las promociones tienen efectos significativamente diferentes sobre las ventas.

EA 11-6

	n	$\bar{x}$	s^2
Sistema A	6	58	10.4000
Sistema B	5	57	38.5000
Sistema C	4	59	35.3333

$$\bar{\bar{x}} = \frac{6(58) + 5(57) + 4(59)}{6 + 5 + 4} = 57.9333$$

$$\hat{\sigma}_b^2 = \frac{\Sigma n_j (x_j - \bar{\bar{x}})^2}{k - 1}$$

$$= \frac{6(58 - 57.9333)^2 + 5(57 - 57.9333)^2 + 4(59 - 57.9333)^2}{3 - 1}$$

$$= \frac{8.9333}{2} = 4.4667$$

$$\hat{\sigma}_w^2 = \sum\left(\frac{n_j - 1}{n_T - k}\right) s_j^2 = \frac{5(10.4) + 4(38.5) + 3(35.3333)}{15 - 3} = \frac{312}{12} = 26$$

$$F = \frac{\hat{\sigma}_w^2}{\hat{\sigma}_w^2} = \frac{4.4667}{26} = 0.17$$

Con 2 grados de libertad en el numerador, 12 grados de libertad en el denominador y $\alpha = 0.05$, el valor crítico de F es 3.89, de manera que no se rechaza H_0 porque 0.17 < 3.89. La efectividad de los sistemas no tiene diferencias significativas.

11.5 Inferencias acerca de una varianza de población

Es necesario tomar decisiones acerca de la variabilidad en una población

En los capítulos 7 a 9, aprendimos cómo formar intervalos de confianza y probar hipótesis de una o dos medias o proporciones de población. Antes en este capítulo, utilizamos las pruebas ji-cuadrada y F para hacer inferencias respecto a más de dos medias o proporciones. Pero no siempre estamos interesados en medias o proporciones. En muchas situaciones, los tomadores de decisiones responsables tienen que hacer inferencias sobre la variabilidad de una población. Con el fin de programar la fuerza de trabajo en la temporada de cosecha, un cultivador de duraznos necesita conocer no sólo el tiempo medio que tardan los duraznos en madurar, sino también su varianza alrededor de ese tiempo medio. Un sociólogo que investiga el efecto de la educación en el poder adquisitivo desea saber si los ingresos de los egresados de la universidad son más variables que los egresados de bachillerato. Los instrumentos de precisión que se utilizan en el trabajo de laboratorio deben ser bastante precisos en promedio; pero, además, las mediciones repetidas deberán mostrar poca variación. En esta sección veremos cómo hacer inferencias acerca de una sola varianza de población; la siguiente trata problemas que implican las varianzas de dos poblaciones.

La distribución de la varianza de la muestra

En respuesta a muchas quejas respecto a la tardanza del correo, el director general del servicio postal inicia una investigación preliminar. Un investigador da seguimiento a nueve cartas desde Nueva York hasta Chicago, para estimar la desviación estándar del tiempo de entrega. La tabla 11-15 da los datos y calcula $\bar{x}$, s^2 y s. Como vimos en el capítulo 7, se usa s para estimar σ.

Determinación de la incertidumbre asociada a las estimaciones de la desviación estándar de la población

Podemos decirle al director general que la desviación estándar de la *población*, estimada a partir de la desviación estándar de la *muestra* es, aproximadamente, 23 horas. Pero también desea saber qué tan precisa es esa estimación y qué incertidumbre se le asocia. En otras palabras, quiere un intervalo de confianza, no nada más una estimación puntual de σ. Para encontrar dicho intervalo, debemos conocer la distribución muestral de s. Es tradicional hablar de s^2 más que de s, pero esto no ocasiona problemas, debido a que siempre podemos pasar de s^2 y de σ^2 a s y σ, obteniendo la raíz cuadrada; y podemos ir en la otra dirección elevando al cuadrado.

Tabla 11-15

Tiempo de entrega (en horas) para las cartas que van de Nueva York a Chicago

Tiempo x	$\bar{x}$	$x - \bar{x}$	$(x - \bar{x})^2$
50	59	− 9	81
45	59	−14	196
27	59	−32	1,024
66	59	7	49
43	59	−16	256
96	59	37	1,369
45	59	−14	196
90	59	31	961
69	59	10	100
$\Sigma x = \mathbf{531}$			$\Sigma(x - \bar{x})^2 = \mathbf{4{,}232}$

$$x = \frac{\Sigma x}{n} = \frac{531}{9} \quad [3\text{-}2]$$

$$= 59 \text{ horas}$$

$$s^2 = \frac{\Sigma(x - \bar{x})^2}{n - 1} = \frac{4{,}232}{8} \quad [3\text{-}17]$$

$$= 529 \text{ horas al cuadrado}$$

$$s = \sqrt{s^2} = \sqrt{529} \quad [3\text{-}18]$$

$$= 23 \text{ horas}$$

Estadístico ji-cuadrada para inferencias sobre una varianza

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} \qquad [11\text{-}12]$$

Si la varianza de la población es σ^2, entonces el estadístico tiene una distribución ji-cuadrada con $n - 1$ grados de libertad. Este resultado es exacto si la población es normal; pero, incluso para muestras tomadas de poblaciones no normales, suele ser una buena aproximación. Podemos ahora usar la distribución ji-cuadrada para formar intervalos de confianza y probar hipótesis respecto a σ^2.

Intervalos de confianza para la varianza de población

Construcción de un intervalo de confianza para una varianza

Suponga que queremos un intervalo de confianza del 95% para el problema de entrega de correo. La figura 11-14 muestra cómo empezar a formarlo.

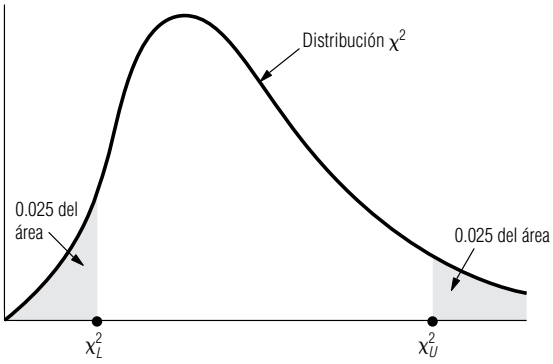


FIGURA 11-14

Construcción de un intervalo de confianza para α^2

Localizamos dos puntos en la distribución χ^2: χ^2_U corta 0.025 del área en la cola superior de la distribución y χ^2_L corta 0.025 del área en la cola inferior de la distribución. (Para un intervalo de confianza del 99%, tendríamos 0.005 del área en cada cola y otras mitades similares para otros niveles de confianza.) Los valores de χ^2_L, y de χ^2_U se pueden encontrar en la tabla 5 del apéndice. En el problema de entrega de correo, con $9 - 1 = 8$ grados de libertad, $\chi^2_L = 2.180$ y $\chi^2_U = 17.535$.

Ahora, la ecuación 11-12 da χ^2 en términos de s^2, n y σ^2. Para obtener un intervalo de confianza para σ^2, despejamos σ^2 de la ecuación 11-12:

Límites superior e inferior para el intervalo de confianza

$$\sigma^2 = \frac{(n-1)s^2}{\chi^2} \qquad [11\text{-}13]$$

y, entonces, nuestro intervalo de confianza está dado por:

Intervalo de confianza para σ^2

$$\sigma^2_L = \frac{(n-1)s^2}{\chi^2_U} \leftarrow \text{Límite inferior de confianza}$$

$$\sigma^2_U = \frac{(n-1)s^2}{\chi^2_L} \leftarrow \text{Límite superior de confianza} \qquad [11\text{-}14]$$

Note que, debido a que χ^2 aparece en el denominador de la ecuación 11-13, podemos usar χ^2_U para encontrar σ^2_L y χ^2_L para encontrar σ^2_U. Continuando con el problema del director general del servicio postal, vemos que puede tener una certeza del 95% de que la varianza de la población está entre 241.35 y 1,941.28 horas al cuadrado:

$$\sigma_L^2 = \frac{(n-1)s^2}{\chi_U^2} = \frac{8(529)}{17.535} = 241.35$$
$$\sigma_L^2 = \frac{(n-1)s^2}{\chi_U^2} = \frac{8(529)}{2.180} = 1{,}941.28 \qquad [11\text{-}14]$$

Así que un intervalo del 95% de confianza para σ sería de $\sqrt{241.35}$ a $\sqrt{1{,}941.28}$ horas, es decir, de 15.54 a 44.06 horas.

Prueba de dos colas de una varianza

Prueba de hipótesis acerca de una varianza: pruebas de dos colas

Un profesor de administración ha diseñado meticulosamente sus exámenes. Con el fin de estar razonablemente seguro de que un examen distingue bien las diferencias en el logro de los estudiantes, la desviación estándar de los resultados del examen no puede ser demasiado pequeña. Por el otro lado, si la desviación estándar es muy grande, habrá una tendencia a tener muchas calificaciones muy bajas, lo cual es deprimente para el ánimo de los estudiantes. Experiencias pasadas han hecho que el profesor crea que una desviación estándar de aproximadamente 13 puntos para un examen de 100, indica que el examen logra un buen equilibrio de estos dos objetivos.

El profesor acaba de hacer un examen a su grupo de 31 estudiantes de primero y segundo año. La calificación promedio fue 72.7 y la desviación estándar de la muestra fue 15.9. ¿Este examen cumple con el criterio del profesor de un buen examen? Podemos resumir los datos:

$\sigma_{H_0} = 13$ ← Valor hipotetizado de la desviación estándar de la población
$s = 15.9$ ← Desviación estándar de la muestra
$n = 31$ ← Tamaño de la muestra

Si el profesor utiliza un nivel de significancia de 0.10 en la prueba de su hipótesis, podemos plantear el problema en símbolos:

Planteamiento del problema en símbolos

H_0: $\sigma = 13$ ← Hipótesis nula: la desviación estándar real es 13 puntos
H_1: $\sigma \neq 13$ ← Hipótesis alternativa: la desviación estándar real no es 13 puntos
$\alpha = 0.10$ ← Nivel de significancia para probar estas hipótesis

Lo primero que se hace es utilizar la ecuación 11-12 para calcular el estadístico χ^2:

Cálculo del estadístico ji-cuadrada

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{30(15.9)^2}{(13)^2} = 44.88 \qquad [11\text{-}12]$$

Interpretación de los resultados

Este estadístico tiene una distribución χ^2 con $n - 1$ (= 30, en este caso) grados de libertad. Aceptaremos la hipótesis nula si ji-cuadrada no es demasiado grande ni demasiado pequeña. De la tabla de la distribución χ^2 (tabla 5 del apéndice), podemos ver que los valores apropiados de χ^2 para que 0.05 del área se encuentre en cada cola de la curva son 18.493 y 43.773. Estos dos límites de la región de aceptación y el estadístico de la muestra observada ($\chi^2 = 44.88$) se ilustran en la figura 11-15. Vemos que el valor muestral de χ^2 no se encuentra en la región de aceptación, de modo que el profesor deberá rechazar la hipótesis nula; el examen no cumple con su criterio de un buen examen.

Prueba de una cola para la varianza

Prueba de hipótesis de la varianza: pruebas de una cola

La empresa Precision Analytics fabrica una amplia línea de instrumentos de precisión y tiene una buena reputación en el mercado por la calidad de sus instrumentos. Con el fin de conservar su repu-

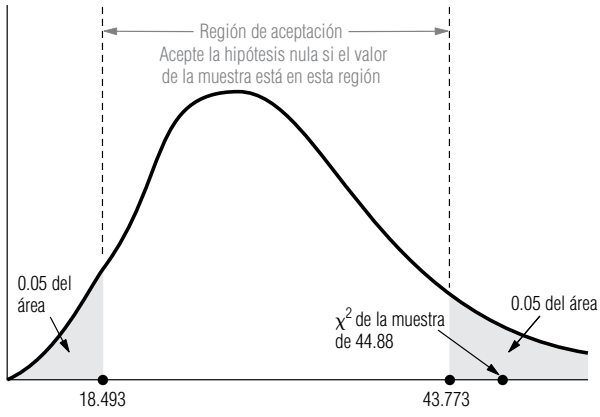


FIGURA 11-15

Prueba de hipótesis de dos colas al nivel 0.10 de significancia; señala la región de aceptación y ji-cuadrada de la muestra

tación, mantiene un estricto control de calidad en todos sus productos. No pone a la venta una balanza analítica, por ejemplo, a menos que muestre una variabilidad significativamente menor que un microgramo (para $\alpha = 0.01$) cuando se pesan cantidades de aproximadamente 500 gramos. La línea de producción acaba de entregar una nueva balanza a la división de control de la calidad.

Se prueba la nueva balanza utilizándola para pesar el mismo peso estándar de 500 gramos 30 veces. La desviación estándar de la muestra fue 0.73 microgramos. ¿Se deberá vender la balanza? Hacemos un resumen de los datos:

$$\sigma_{H_0} = 1 \quad \leftarrow \text{Valor hipotético de la desviación estándar de la población}$$
$$s = 0.73 \leftarrow \text{Desviación estándar de la muestra}$$
$$n = 30 \quad \leftarrow \text{Tamaño de la muestra}$$

y planteamos el problema:

Planteamiento del problema en símbolos

$$H_0: \sigma = 1 \quad \leftarrow \text{Hipótesis nula: la desviación estándar verdadera es 1 microgramo}$$
$$H_1: \sigma < 1 \quad \leftarrow \text{Hipótesis alternativa: la desviación estándar verdadera es menor que 1 microgramo}$$
$$\alpha = 0.01 \leftarrow \text{Nivel de significancia para probar estas hipótesis}$$

Cálculo del estadístico χ^2

Empezamos por utilizar la ecuación 11-12 para calcular el estadístico χ^2:

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} \qquad [11\text{-}12]$$
$$= \frac{29(0.73)^2}{(1)^2}$$
$$= 15.45$$

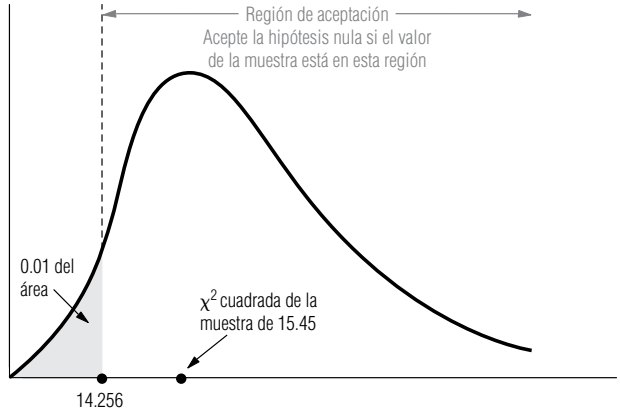


FIGURA 11-16

Prueba de hipótesis de una cola al nivel de significancia de 0.01; indica la región de aceptación y la χ^2 de la muestra

Interpretación de los resultados

Rechazamos la hipótesis nula y enviamos la balanza a ventas si este estadístico es lo suficientemente pequeño. De la tabla 5 del apéndice, vemos que con 29 grados de libertad (30 − l), el valor χ^2 que deja un área de 0.01 en la cola inferior de la curva es 14.256. La región de aceptación y el valor observado de ji-cuadrada se ilustran en la figura 11-16. Vemos que no podemos rechazar la hipótesis nula. La balanza se tendrá que regresar a la línea de producción para ajustes.

SUGERENCIAS Y SUPOSICIONES

Hasta ahora, se ha visto cómo hacer inferencias acerca de una, dos o varias *medias* o *proporciones*. Pero también es interesante hacer inferencias acerca de la variabilidad de la población. Para una población, esto se logra con la varianza de la muestra y la distribución ji-cuadrada. Advertencia: la prueba de ji-cuadrada puede ser de una o dos colas. Sugerencia: si la pregunta que debe contestarse incluye las palabras *menor que*, *mayor que*, *menor o igual que* o *mayor o igual que*, se usa la prueba de una cola; si la pregunta se refiere a diferente de o cambio de, se usa la prueba de dos colas.

Ejercicios 11.5

Ejercicios de autoevaluación

EA 11-7 Dado que 127 es la varianza de la muestra para un conjunto de 9 observaciones, construya un intervalo de confianza del 95% para la varianza de la población.

EA 11-8 Un gerente de producción siente que la tasa de producción de los empleados con experiencia es seguramente mayor que la de los nuevos empleados, pero no espera que la variabilidad en las tasas de producción difiera entre los dos grupos. En estudios anteriores se ha encontrado que la producción promedio por hora para los nuevos empleados en este tipo de trabajo específico es 20 unidades por hora con una varianza de 56 unidades al cuadrado. Para un grupo de 20 empleados con 5 años de experiencia, la producción promedio en este mismo tipo de trabajo es 30 unidades por hora con varianza muestral de 28 unidades al cuadrado. ¿Parecería que la variabilidad en la producción difiere entre los dos niveles de experiencia? Pruebe las hipótesis para un nivel de significancia de 0.05.

Conceptos básicos

■ **11-39** Una muestra de 20 observaciones de una distribución normal tiene media de 37 y varianza de 12.2. Construya un intervalo de confianza del 90% para la varianza verdadera de la población.

■ **11-40** Se tiene la hipótesis de que la desviación estándar de una distribución es 50. Si una muestra observada de 30 elementos produce una desviación estándar de muestra de 57, ¿deberemos rechazar la hipótesis nula de que la desviación estándar verdadera es 50? Utilice un nivel de significancia de 0.05.

■ **11-41** Dada una desviación estándar muestral de 6.4 para una muestra de 15 observaciones, construya un intervalo de confianza del 90% para la varianza de la población.

Aplicaciones

■ **11-42** Un fabricante de telescopios desea que sus aparatos tengan desviaciones estándar en su resolución que sean significativamente menores que 2 al enfocar un objeto a una distancia de 500 años luz. Cuando se usa un nuevo telescopio para enfocar un objeto a una distancia de 500 años luz 30 veces, la desviación estándar de la muestra resulta ser 1.46. ¿Debe el fabricante vender este telescopio?

a) Establezca las hipótesis nula y alternativa explícitas.
b) Pruebe las hipótesis para $\alpha = 0.01$.
c) Establezca una conclusión explícita.

■ **11-43** MacroSwift diseñó un nuevo sistema operativo que revolucionará la industria de la computación. El único problema es que la compañía espera que el tiempo promedio requerido para aprender a manejarlo sea 124 horas. Aun cuando este tiempo de aprendizaje es largo, la compañía está en verdad preocupada por la varianza de ese tiempo. Los datos preliminares indican que la varianza es 171 horas al cuadrado. En pruebas recientes con 25 personas se encontró un tiempo promedio de aprendizaje de 123 horas y una va-

rianza de la muestra de 196.5 horas al cuadrado. ¿Indican estos datos que la variabilidad en el tiempo de aprendizaje es diferente de la estimación previa? Pruebe sus hipótesis a un nivel de significancia de 0.02.

■ **11-44** Un sicólogo está al tanto de los estudios que indican que la variabilidad en lapso de atención de los pacientes de cinco años de edad se puede resumir por $\sigma^2 = 64$ minutos al cuadrado. Se pregunta si el lapso de atención de pacientes de seis años es diferente. Una muestra de 20 pacientes de seis años da $s^2 = 28$ minutos al cuadrado.

a) Establezca explícitamente las hipótesis nula y alternativa.
b) Pruebe sus hipótesis para $\alpha = 0.05$.
c) Establezca una conclusión explícita.

■ **11-45** Al verificar sus automóviles para saber si cumplen con las normas de emisión de contaminantes establecidas por el gobierno, un fabricante de automóviles midió la emisión de 30 vehículos. Encontró que el número promedio de partículas de contaminantes emitidas estaba dentro de los niveles requeridos, pero la varianza de la muestra fue 50. Encuentre un intervalo de confianza del 90% para la varianza en la emisión de partículas para estos automóviles.

■ **11-46** Un banco está considerando estrategias de reducción de costos asociados con las cuentas de ahorro. Ha encontrado que la varianza en el número de días entre transacciones para esas cuentas es 80 días al cuadrado. El banco desea reducir la varianza desalentando el uso de las cuentas para guardar dinero por poco tiempo. Por tanto, después de implantar una nueva política que penaliza al cliente con cargos de servicio por más de un retiro al mes, el banco decide probar si hubo un cambio en la varianza de los días entre transacciones. Para una muestra de 25 cuentas de ahorro, se encuentra que la varianza entre transacciones es 28 días al cuadrado. ¿Está justificado el banco al asegurar que la nueva política reduce la varianza de días entre transacciones? Pruebe las hipótesis al nivel de significancia de 0.05.

■ **11-47** Sam Bogart, dueño de la compañía de aparatos de audio, Play-It-Again, ofrece un año de garantía en todos los estéreos que vende su compañía. Para los 30 aparatos a los que se les dio servicio de garantía el año anterior, el costo promedio de compostura fue \$75 y la desviación estándar de la muestra fue \$15. Calcule un intervalo de confianza del 95% para la desviación estándar real del costo de reparación. Sam ha decidido que a menos que la desviación estándar verdadera sea menor que \$20, comprará sus aparatos de audio a otro mayorista. Ayude a Sam a probar las hipótesis adecuadas, utilizando un nivel de significancia de 0.01. ¿Debe cambiar de proveedor?

Soluciones a los ejercicios de autoevaluación

EA 11-7 Para un intervalo de confianza del 95% con 8 grados de libertad:

$$\sigma_L^2 = \frac{(n-1)s^2}{\chi_U^2} = \frac{8(127)}{17.535} = 57.941$$

$$\sigma_U^2 = \frac{(n-1)s^2}{\chi_L^2} = \frac{8(127)}{2.180} = 466.055$$

Así, el intervalo de confianza es (57.941, 466.055).

EA 11-8 Para probar H_0: $\sigma^2 = 56$ contra H_1: $\sigma^2 \neq 56$ para $\alpha = 0.05$, los límites de la región de aceptación son

$$\chi^2 = 8.907 \quad \text{y} \quad \chi^2 = 32.852$$

El valor observado es $\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{19(28)}{56} = 9.5$, y no se rechaza H_0; la variabilidad no es significativamente diferente.

11.6 Inferencias acerca de las varianzas de dos poblaciones

Comparación de las varianzas de dos poblaciones

En el capítulo 9 vimos varias situaciones en las que deseábamos comparar las medias de dos poblaciones distintas. Recuerde que hicimos esto buscando la *diferencia* de las medias de dos muestras tomadas de dichas poblaciones. Aquí, queremos comparar las varianzas de dos poblaciones. Sin em-

bargo, más que buscar la *diferencia* de dos varianzas muestrales, es más conveniente estudiar su cociente. Los siguientes dos ejemplos muestran cómo se hace esto.

Prueba de una cola de dos varianzas

Una prominente socióloga de una importante universidad del medio oeste estadounidense cree de que los ingresos de los graduados de la universidad tienen una variabilidad mucho mayor que los ingresos de las personas que no cursaron la universidad. Con el fin de probar esta teoría, envía a dos ayudantes de investigación a Chicago a investigar los ingresos de estas dos poblaciones. El primer ayudante toma una muestra aleatoria de 21 graduados de la universidad y encuentra que sus ingresos tienen una desviación estándar de la muestra $s_1 = \$17{,}000$. El segundo ayudante toma una muestra de 25 no graduados y obtiene una desviación estándar en los ingresos $s_2 = \$7{,}500$. Los datos de nuestro problema se pueden resumir de la siguiente manera:

Datos del problema

$$s_1 = 17{,}000 \leftarrow \text{Desviación estándar de la primera muestra}$$
$$n_1 = 21 \leftarrow \text{Tamaño de la primera muestra}$$
$$s_2 = 7{,}500 \leftarrow \text{Desviación estándar de la segunda muestra}$$
$$n_2 = 25 \leftarrow \text{Tamaño de la segunda muestra}$$

Por qué es apropiada la prueba de una cola

Debido a que la socióloga tiene la teoría de que los ingresos de los egresados de la universidad son *más* variables que los ingresos de las personas que no cursaron la universidad, una prueba de una cola es apropiada. La socióloga desea verificar su teoría al nivel de significancia de 0.01. Podemos establecer de manera formal sus hipótesis:

Planteamiento de las hipótesis

$$H_0: \sigma_1^2 = \sigma_2^2 \ (\text{o } \sigma_1^2/\sigma_2^2 = 1) \leftarrow \text{Hipótesis nula: las dos varianzas son iguales}$$
$$H_1: \sigma_1^2 > \sigma_2^2 \ (\text{o } \sigma_1^2/\sigma_2^2 > 1) \leftarrow \text{Hipótesis alternativa: los ingresos de los egresados de la universidad tienen una varianza más grande}$$
$$\alpha = 0.01 \leftarrow \text{Nivel de significancia para probar estas hipótesis}$$

Sabemos que s_1^2 se puede utilizar para estimar σ_1^2, y s_2^2 se puede usar para estimar σ_2^2. Si la hipótesis alternativa es verdadera, esperaríamos que s_1^2 fuera mayor que s_2^2 (o de manera equivalente que s_1^2/s_2^2 fuera mayor que uno). Pero, ¿qué tanto deberá ser mayor s_2^1 para que podamos rechazar la hipótesis nula? Para responder a esta pregunta, debemos conocer la distribución de s_2^1/s_2^2. Si suponemos que las dos poblaciones están razonablemente bien descritas por distribuciones normales, entonces el cociente:

Descripción del estadístico *F*

Cociente *F* para inferencias acerca de dos varianzas

$$F = \frac{s_1^2}{s_2^2} \qquad [11\text{-}15]$$

tiene una distribución F con $n_1 - 1$ grados de libertad en el numerador y $n_2 - 1$ grados de libertad en el denominador.

En el problema de los ingresos calculamos el estadístico F de la muestra:

$$F = \frac{s_1^2}{s_2^2} \qquad [11\text{-}15]$$
$$= \frac{(17{,}000)^2}{(7{,}500)^2}$$
$$= \frac{289{,}000{,}000}{56{,}250{,}000}$$
$$= 5.14$$

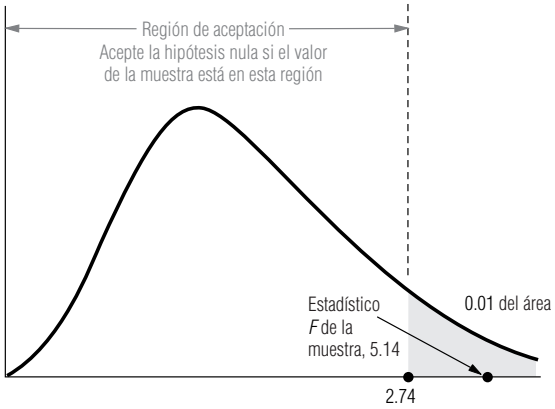


FIGURA 11-17

Prueba de hipótesis de una cola al nivel de significancia de 0.01; indica la región de aceptación y el estadístico *F*

Interpretación de los resultados

Para 20 grados de libertad (21 − 1) en el numerador y 24 grados de libertad (25 − 1) en el denominador, la tabla 6 del apéndice nos indica que el valor crítico que separa a las regiones de aceptación y de rechazo es 2.74. La figura 11-17 muestra la región de aceptación y el estadístico *F* observado de 5.14. Nuestra socióloga rechaza la hipótesis nula y concluye que los datos de la muestra apoyan su teoría.

Manejo de pruebas de cola inferior en la tabla 6 del apéndice

En este punto son necesarias unas palabras de advertencia respecto al uso de la tabla 6 del apéndice. Se habrá dado cuenta de que la tabla da los valores del estadístico *F* que son apropiados sólo para pruebas de *cola superior*. ¿De qué manera podemos manejar hipótesis alternativas de la forma $\sigma_1^2 < \sigma_2^2$ (o $\sigma_1^2/\sigma_2^2 <$ l)? Es sencillo si notamos que $\sigma_1^2/\sigma_2^2 < 1$ es equivalente a $\sigma_2^2/\sigma_1^2 > 1$. Así, todo lo que necesitamos es calcular el cociente s_2^2/s_1^2, que también sigue una distribución *F* (pero con $n_2 - 1$ grados de libertad en el numerador y $n_1 - 1$ grados de libertad en el denominador), y después usamos la tabla 6 del apéndice. Existe otra forma de decir lo mismo: **siempre que realice una prueba de una cola de dos varianzas, numere las poblaciones de modo que la hipótesis alternativa tenga la forma:**

$$H_0: \sigma_1^2 > \sigma_2^2 \text{ (o } \sigma_1^2/\sigma_2^2 > 1)$$

y luego proceda como en el ejemplo de los ingresos.

Una prueba de dos colas de dos varianzas

Búsqueda del valor crítico en una prueba de dos colas

El procedimiento para la prueba de dos colas de dos varianzas es parecido al de la prueba de una cola. El único problema surge en la búsqueda del valor crítico de la cola inferior. Éste se relaciona con el problema de las pruebas de cola inferior analizadas en el último párrafo, y lo resolveremos de manera parecida.

Un criterio para la evaluación de anestésicos orales odontológicos es la variabilidad de la cantidad de tiempo entre la inyección y la pérdida completa de sensibilidad en el paciente. (Esto se conoce como tiempo de demora del efecto.) Una compañía farmacéutica grande ha desarrollado dos nuevos anestésicos orales, que comercializará con los nombres de Oralcaine y Novasthetic. A partir de similitudes en la estructura química de los dos compuestos, se ha predicho que deben mostrar la misma varianza en el tiempo de demora del efecto. En la tabla 11-16 se presentan los datos para las pruebas sobre los dos compuestos (en las que se controlaron otras variables como la edad y el peso del paciente).

Tabla 11-16

Tiempos de demora del efecto de dos anestésicos

Anestésico	Tamaño de muestra (n)	Varianza de muestra (segundos al cuadrado) (s^2)
Oralcaine	31	1,296
Novasthetic	41	784

La compañía desea probar a un nivel de significancia del 2% si los dos anestésicos tienen la misma varianza en el tiempo de demora del efecto. En símbolos, las hipótesis son:

Planteamiento de las hipótesis

$$H_0: \sigma_1^2 = \sigma_2^2 \ (\text{o } \sigma_1^2/\sigma_2^2 = 1) \quad \leftarrow \text{Hipótesis nula: las dos varianzas son iguales}$$
$$H_1: \sigma_1^2 \neq \sigma_2^2 \ (\text{o } \sigma_1^2/\sigma_2^2 \neq 1) \quad \leftarrow \text{Hipótesis alternativa: las dos varianzas son diferentes}$$
$$\alpha = 0.02 \leftarrow \text{Nivel de significancia de la prueba}$$

Cálculo del estadístico *F*

Para probar estas hipótesis, de nuevo hacemos uso de la ecuación 11-15:

$$F = \frac{s_1^2}{s_2^2} \qquad [11\text{-}15]$$
$$= \frac{1{,}296}{784}$$
$$= 1.65$$

Este estadístico viene de una distribución F con $n_1 - 1$ grados de libertad en el numerador (30, en este caso) y $n_2 - 1$ grados de libertad en el denominador (40, en este caso). Usemos la notación:

Notación útil para la prueba

$$F(n, d, \alpha)$$

para denotar el valor de F con n grados de libertad en el numerador, d grados de libertad en el denominador y un área α en la cola superior. En nuestro problema, la región de aceptación se extiende de $F(30, 40, 0.99)$ a $F(30, 40, 0.01)$, como se ilustra en la figura 11-18.

Podemos obtener el valor de $F(30, 40, 0.01)$ directamente de la tabla 6 del apéndice; éste es 2.20. Sin embargo, el valor de $F(30, 40, 0.99)$ no está en la tabla. Ahora bien, $F(30, 40, 0.99)$ corresponderá a un valor *pequeño* de s_1^2/s_2^2 y, por tanto, a un valor *grande* de s_2^2/s_1^2, que es justo el recíproco de s_1^2/s_2^2. Dado el análisis de las pruebas de cola inferior que acabamos de hacer, podríamos sospechar que:

Valor de cola inferior de *F* para pruebas de dos colas

$$F(n, d, \alpha) = \frac{1}{F(d, n, 1 - \alpha)} \qquad [11\text{-}16]$$

y esto resulta ser verdadero. Podemos utilizar esta ecuación para encontrar $F(30, 40, 0.99)$:

$$F(30, 40, 0.99) = \frac{1}{F(40, 30, 0.01)}$$
$$= \frac{1}{2.30}$$
$$= 0.43$$

Interpretación de los resultados

En la figura 11-19 se ilustró la región de aceptación para esta prueba de hipótesis y el valor observado de F. Vemos ahí que la hipótesis nula se acepta, de modo que concluimos que la diferencia ob-

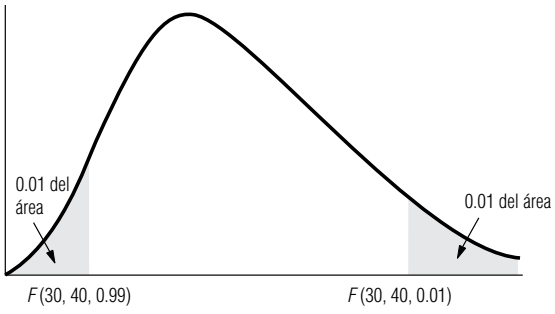


FIGURA 11-18

Prueba de hipótesis de dos colas al nivel de significancia de 0.02

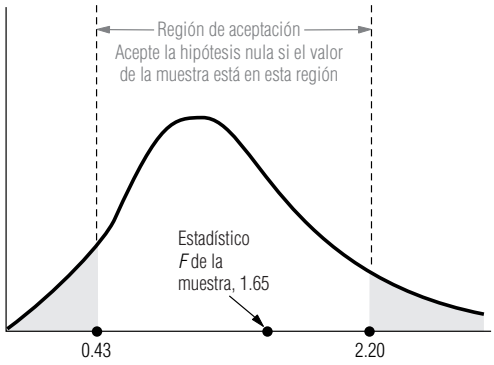


FIGURA 11-19

Prueba de hipótesis de dos colas al nivel de significancia de 0.02; indica la región de aceptación y el estadístico *F* de la muestra

servada en las varianzas de muestras de los tiempos de demora del efecto para los dos anestésicos no es estadísticamente significativa.

SUGERENCIAS Y SUPOSICIONES

Esta sección se refiere al uso de la prueba *F* para comparar las varianzas de dos *poblaciones* examinando el cociente de las varianzas de dos *muestras*. Adevertencia: la tabla 6 del apéndice proporciona valores de *F* que son apropiados sólo para las pruebas de la *cola superior*. Sugerencia: si desea hacer una prueba de *cola inferior*, asegúrese de convertirla a una prueba de cola superior como se hizo un par de páginas atrás. Si desea hacer una prueba de dos colas, utilice la ecuación 11-16 para convertir un valor de cola superior de la tabla en el de cola inferior requerido para la prueba.

Ejercicios 11.6

Ejercicios de autoevaluación

EA 11-9 Un supervisor de control de calidad de una empresa fabricante de automóviles está preocupado por la uniformidad del número de defectos en los automóviles que salen de la línea de ensamble. Si una línea de ensamble tiene una variabilidad significativamente mayor en el número de defectos, entonces es necesario hacer cambios. El supervisor reunió los siguientes datos:

	Número de defectos	
	Línea de ensamble A	**Línea de ensamble B**
Media	10	11
Varianza	9	25
Tamaño de muestra	20	16

¿La línea de ensamble B tiene una variabilidad significativamente mayor en el número de defectos? Pruebe al nivel de significancia de 0.05.

EA 11-10 Techgene, Inc. está preocupada por la variabilidad en el número de bacterias producidas por distintos cultivos. Si los cultivos tienen una variabilidad significativamente diferente, entonces se crea confusión en los experimentos y se producen resultados extraños (se entiende que la directiva de la compañía se ponga ansiosa cuando los científicos comienzan a producir cosas extrañas). Se ha recopilado la siguiente información:

	Número de bacterias (en miles)										
Cultivo A	91	89	83	101	93	98	144	118	108	125	138
Cultivo B	62	76	90	75	88	99	110	140	145	130	110

a) Calcule s_A^2 y s_B^2.
b) Establezca las hipótesis nula y alternativa explícitas, y pruebe al nivel de significancia de 0.02.

Conceptos básicos

■ **11-48** Para dos poblaciones que se piensa tienen la misma varianza, se tiene la siguiente información. Una muestra de 16 tomada de la población 1 exhibe una varianza muestral de 3.75, y una muestra de 10 tomada de la población 2 exhibe una varianza de 5.38.

a) Calcule el cociente *F* para la prueba de igualdad de varianzas.
b) Encuentre el valor *F* crítico para la cola superior, con un nivel de significancia de 0.10.
c) Encuentre el valor *F* correspondiente a la cola inferior.
d) Establezca la conclusión de su prueba.

■ **11-49** En un estudio de comparación entre las medias de dos grupos, se observó que la forma más común de la prueba *t* de dos grupos para la diferencia entre dos medias supone que las varianzas de población para los dos grupos son iguales. Un experimentador, usando una condición de control y una condición experimental en su estudio de reacción a la droga, desea verificar que esta suposición se cumple, es decir, que el tratamiento administrado afecta sólo a la media, y no a la varianza de la variable que se investiga. A partir de estos datos, calculó la varianza del grupo experimental en 25.8 y la del grupo de control en 20.6. El grupo experimental tenía 25 sujetos, mientras que el del grupo de control, 31. ¿Puede el experimentador proceder a usar la prueba *t*, que supone varianzas iguales para los dos grupos? Utilice $\alpha = 0.10$.

■ **11-50** De una muestra de 25 observaciones, la estimación de la desviación estándar de la población fue 15.0. Para otra muestra de 14 observaciones, la estimación fue 9.7. ¿Podemos aceptar la hipótesis de que las dos muestras provienen de poblaciones con varianzas iguales o debemos concluir que la varianza de la segunda población es menor? Utilice el nivel de significancia de 0.01.

Aplicaciones

■ **11-51** El señor Raj, un inversionista, ha reducido su búsqueda de un fondo mutualista a los fondos Oppy y MLPFS. La tasa de rendimiento de Oppy es más baja, pero parece más estable que la de MLPFS. Si la variabilidad en la tasa de rendimiento de Oppy es, en efecto, significativamente más baja que la de MLPFS, entonces invertirá su dinero en la primera. Si no hay una diferencia significativa en la variabilidad, hará la inversión en MLPFS. Para tomar una decisión, Raj ha obtenido una muestra de 21 tasas mensuales de rendimiento de ambas empresas. Para Oppy la desviación estándar fue 2, y para MLPFS fue 3. ¿En qué empresa invertirá su dinero el señor Raj? Pruebe para $\alpha = 0.05$.

■ **11-52** Una compañía de seguros está interesada en la duración de la hospitalización por varias enfermedades. Seleccionaron una muestra aleatoria de 20 pacientes del hospital A y 25 del hospital B internados por la misma enfermedad. El tiempo que pasaron en el hospital A tuvo un promedio de 2.4 días con desviación estándar de 0.6 días. El tiempo de tratamiento en el hospital B promedió 2.3 días con desviación estándar de 0.9 días. ¿Los pacientes del hospital A tiene significativamente menor variabilidad en su tiempo de recuperación? Pruebe para un nivel de significancia de 0.01.

■ **11-53** La Nation's Broadcasting Company está interesada en el número de personas que sintonizan sus programas de éxito *Buddies* y *Ride to Nowhere*; más importante, la compañía está muy preocupada por la variabilidad en el número de personas que ven los programas. Los anunciantes quieren televidentes consistentes con la esperanza de que la exposición constante a los anuncios ayude a vender sus productos. Los datos son los siguientes (en millones de televidentes) para los últimos meses.

	Número de televidentes (millones)									
Buddies	57.4	62.6	54.6	52.4	60.5	61.8	71.4	67.5	62.6	58.4
Ride to Nowhere	64.5	58.2	39.5	24.7	40.2	41.6	38.4	33.6	34.4	37.8

a) Calcule $s^2_{BUDDIES}$ y s^2_{RIDE}.
b) Establezca las hipótesis explícitas para determinar si la variabilidad es la misma entre las dos poblaciones. Pruebe para 0.10 de nivel de significancia.

■ **11-54** La HAL Corporation está a punto de poner en el mercado una computadora personal nueva más rápida, PAL, que sustituirá a su viejo modelo, CAL. Aunque, en promedio, PAL es más rápida que CAL, la velocidad de procesamiento de la nueva computadora parece más variable. (La velocidad de procesamiento depende del programa que se esté corriendo, la cantidad de datos de entrada y la cantidad de datos de sa-

lida.) Se suministraron a PAL y CAL dos muestras de 25 corridas que cubrían la variedad de tareas esperadas (una muestra a cada una). Los resultados fueron los siguientes:

	Tiempo de procesamiento (en centésimas de segundo)	
	PAL	CAL
Media	50	75
Desviación estándar	20	10

Al nivel de significancia de 0.05, ¿es la velocidad de procesamiento de PAL significativamente más variable que la de CAL?

■ **11-55** Dos gerentes de marca estaban en desacuerdo respecto a si las amas de casas urbanas tienen mayor variabilidad en sus patrones de compra de abarrotes que las amas de casa rurales. Para probar sus ideas contrapuestas, tomaron muestras aleatorias de 70 amas de casa de áreas urbanas y 60 de áreas rurales. Encontraron que la varianza en días al cuadrado entre salidas de compras para las amas de casa urbanas fue 14, y la varianza de la muestra para las amas de casa rurales fue 3.5. ¿Es significativa la diferencia de las varianzas en días entre salidas de compras al nivel 0.0l?

■ **11-56** Dos tiendas de helado competidores, Yum-Yum y Goody, anuncian la venta de bolas de helado de 1/4 de libra. Existe cierta preocupación por la variabilidad del tamaño de las bolas, de modo que dos miembros de un grupo de consumidores locales muestrearon 25 raciones de helado de la Yum-Yum y 11 raciones de helado de la Goody. Desde luego, ambos miembros ahora sufren dolor de estómago, de manera que debe ayudarles. ¿Existe alguna diferencia en la varianza del peso de las raciones de helado entre la Yum-Yum y la Goody? Se han recolectado los datos siguientes. Pruebe al nivel 0.10.

	Peso por ración (en centésimos de libra)	
	Yum-Yum	Goody
Media	25	25
Varianza	16	10

Soluciones a los ejercicios de autoevaluación

EA 11-9 H_0: $\sigma_B^2 = \sigma_A^2$

H_1: $\sigma_B^2 > \sigma_A^2$

$$F \text{ observada} = \frac{s_B^2}{s_A^2} = \frac{25}{9} = 2.778$$

$$F_{CRIT} = F_{0.05}\,(15, 19) = 2.23$$

Entonces, se rechaza H_0; la línea de ensamble B tiene una variabilidad significativamente mayor en el número de defectos, por lo que deberán hacerse algunos cambios. (Nota: sólo se está verificando la uniformidad; los autos pueden estar uniformemente mal.)

EA 11-10 a) $s_A^2 = 423.4 \qquad s_B^2 = 755.818$

b) H_0: $\sigma_A^2 = \sigma_B^2$

H_1: $\sigma_A^2 \neq \sigma_B^2$

$$F \text{ observada} = \frac{s_B^2}{s_A^2} = \frac{423.4}{755.818} = 0.56$$

$$F_{0.01}(10, 10) = 4.85$$

$$F_{0.99}(10, 10) = \frac{1}{F\,(10, 10, 0.01)} = \frac{1}{4.85} = 0.21$$

Entonces, se acepta H_0; la administración no tiene que preocuparse por cosas extrañas en el laboratorio.

Estadística en el trabajo

Loveland Computers

Caso 11: ji-cuadrada y ANOVA Tom Hodges ha sido supervisor del equipo de apoyo técnico de la empresa Loveland Computers durante un poco más de un año. Como muchas proveedoras de computadoras, Loveland contrató a una compañía de servicios a nivel nacional para proporcionar un año de reparaciones a domicilio. Esta garantía fue importante para inducir a los clientes a comprar computadoras por teléfono. Pero Loveland ha encontrado que más del 90% de los problemas de los clientes podría haberse solucionado con leer el manual de instrucciones incluido con cada máquina, y el 95% de todos los problemas podría haberse resuelto con instrucciones por teléfono si se animara a los clientes a llamar la atención al cliente antes de pedir la reparación a domicilio. Para ahorrar en costos de garantía, la Loveland invirtió fuertes sumas en su centro de apoyo al cliente, en donde 24 técnicos responderían a las llamadas.

El personal de apoyo al cliente era de dos tipos. La mayoría no tenía mucha experiencia con computadoras. Este personal de apoyo en primer nivel fue reclutado por sus habilidades en el trato telefónico y se capacitó internamente para verificar una lista de rutina de los problemas más comunes. Cuando no podían corregir el problema de un cliente con el protocolo estándar, o cuando un cliente llamaba con una pregunta "difícil", la llamada se transfería a un técnico. Algunos técnicos eran empleados de tiempo completo, pero Hodges había descubierto que podía encontrar suficiente ayuda de tiempo parcial entre los estudiantes de las carreras de ciencias de la computación e ingeniería de la universidad local. Para ajustarse al horario de sus clases, la mayoría trabajaba en un turno que empezaba a las 4:00 p.m.

Entre los problemas que manejaba el personal de primer nivel se incluía ayudar a los clientes a cargar programas al disco duro desde discos flexibles y a verificar las conexiones de los cables. Los técnicos manejaban problemas como la incompatibilidad de algunos programas "residentes en memoria" y cómo recuperar datos "perdidos".

Los directores de varios departamentos se reunían para planear una estrategia con el fin de mejorar el apoyo por teléfono. La clasificación del servicio de Loveland ha bajado de "excelente" a "bueno", según una encuesta de realizada por una compañía de investigación de mercados. Walter Azko pidió a Lee que "asistiera a la reunión para ver si podía ayudar".

Margot Derby, directora de comercialización, empezó la junta con aire de determinación: "Tom, el problema es obvio. Cuando llamamos a las personas que nos enviaron cartas de reclamación, dicen que nunca pueden comunicarse con un técnico. Hablan con el personal de apoyo de primer nivel y luego esperan para siempre. Es obvio que son los clientes de las empresas grandes quienes tienen preguntas 'difíciles' que no puede contestar el personal de primer nivel. Lo único que tienes que hacer es programar más técnicos en los turnos matutinos."

Hodges respondió, "por el contrario, Margot. Son los usuarios domésticos los que necesitan hablar con los técnicos, de manera que la mayoría de las llamadas llega en el turno de la tarde. Nos hacen preguntas 'científicas' mientras juegan con sus máquinas al regresar del trabajo. En cualquier caso, los técnicos están ocupados en el turno de la tarde; obtendré un informe impreso de su tiempo total en el teléfono."

"Sí, pero te apuesto a que si observas el tiempo promedio de llamadas, es más alto por las tardes. Creo que tus técnicos se ponen a platicar con los clientes para llenar el tiempo."

"Bueno, es claro que necesitamos saber cuándo entran las llamadas 'difíciles'", intervino Lee con la esperanza de que la discusión tuviera una dirección más productiva. "Como nadie habla con un técnico sin antes hablar con alguien del personal de primer nivel, podemos hacer que este personal clasifique cada pregunta como fácil o difícil para reunir datos de cada turno. Luego podemos hacer una prueba para ver si realmente hay más preguntas técnicas en la mañana o en la tarde."

"No olvides que son mis clientes de empresas quienes tienen más preguntas difíciles", dijo Margot.

"Todavía creo que no tienes razón. Y a propósito, tengo el presentimiento de que los días de la semana influyen en que las cosas sean diferentes", añadió Tom. "Recibimos un montón de llamadas técnicas al inicio de la semana, pero se reducen al final."

Preguntas de estudio: ¿En qué formato deberán tabularse los datos? ¿Qué prueba estadística podría ser útil si Lee enfoca su atención en la cuestión de los turnos (hace a un lado los comentarios de los clientes corporativos y el día de la semana)? ¿Y qué técnica sería más útil para examinar los efectos del tipo de cliente, turno y día de la semana? ¿Qué podría distorsionar los datos que Lee pide al grupo de apoyo al cliente que recolecte?

Ejercicio de base de datos computacional

HH Industries

Stan Hutchings, director de ventas, metió la cabeza en la oficina de Laurel una mañana, un par de semanas después de la reunión de ventas. "¿Tienes unos minutos? Quiero preguntarte sobre cierto asunto."

"Claro", respondió Laurel, haciendo a un lado una pila de papeles con datos. "¿Qué puedo hacer por ti?"

"Desde hace algún tiempo estamos jugando con la idea de instituir un programa de comisiones por 'ventas desde adentro'. Como están las cosas, sólo nuestros dos representantes externos de ventas trabajan con un salario más comisiones.

Sin embargo, la gente que trabaja en los teléfonos dentro de la compañía son los responsables de casi el 75% de las ventas totales, y han estado pidiendo un pequeño incentivo en forma de comisión. Salió a colación en la junta directiva pasada, y se propusieron algunas buenas ideas. Debo mencionar que alguna vez intentamos poner en práctica un programa de este tipo, pero no funcionó. El truco es que las llamadas se reciben aleatoriamente y las contesta el representante de ventas que en ese momento esté libre. Algunas personas sólo llaman para pedir información y otras para hacer un pedido. Como suele suceder, a menudo un cliente llama más de una vez y, tal vez habla con varios representantes, el resultado es una complicación al asignar la comisión a una sola persona. La solución que propusieron es en realidad innovadora.

"Como ya lo sabes, cada mes nos fijamos un objetivo de ventas", continuó Stan. "El plan propuesto implica mantener las cosas como están hasta que lleguemos al 95% de nuestro objetivo mensual. Después de eso, se agrupará un porcentaje del dinero de las ventas y se repartirá por igual al final del mes. Nuestro propósito es propiciar un espíritu de equipo, eliminando la competencia malsana que suelen propiciar los programas de comisiones."

"Suena bien", comentó Laurel. "Pero, ¿cómo puedo ayudar?"

"Bueno, en aras de la justicia, necesitamos estar seguros de que las ventas realmente siguen un comportamiento aleatorio. Es decir, si mi suposición es correcta, el promedio diario de ventas de cada representante deberá ser el mismo a la larga."

"Ah", exclamó Laurel. "Ya veo qué necesitas. Si quieres conseguirme algunos datos para los diferentes vendedores internos, podría verificar algunos datos para ver cuál es la posibilidad de que en realidad vengan de la misma distribución."

"Lotería. Pensé que podrías ayudarnos. Te voy a conseguir los datos ahora mismo. Gracias, Laurel", Stan se dirigió a la puerta de la oficina.

En los archivos CH11.xxx del CD que acompaña al libro se encuentran los datos de seis meses de ventas diarias de los cuatro representantes internos de tiempo completo.

1. ¿Los datos muestreados provienen de poblaciones con la misma media? Pruebe esta suposición al nivel de significancia del 1%.

Stan observó los resultados. "Hmmm, parece que Mike está solo en una categoría. Bueno, tiene más tiempo en esto que los demás y hay ciertos clientes que preguntan específicamente por él. ¿Qué tal si trabajamos un programa de comisiones para él solo y echamos a andar la propuesta para los otros tres? ¿Se vale hacerlo?"

"Tendré que hacer otra corrida para estar segura", respondió Laurel, "pero creo que sería mejor de esa forma".

2. Verifique la conclusión de Stan respecto a Mike, observando las medias y las desviaciones estándar por vendedor. Pruebe si los datos correspondientes a Debbie, Jeff y Barry parecen indicar poblaciones con la misma media. (Utilice el nivel de significancia del 1%.)

Más tarde, con unos cuantos minutos de tiempo libre, Laurel decidió regresar y verificar sus suposiciones sobre el estudio del personal que había realizado para Stan (vea el capítulo 5). Ella sabía que las recomendaciones se basaban con mucho en suponer una distribución de Poisson para la llegada de las llamadas.

3. Verifique los datos de las llamadas telefónicas de los archivos CH05.xxx del CD que acompaña al libro para ver si parecen seguir una distribución de Poisson. Utilice los intervalos 0-20, 21-25, 26-30, 31-35, $\geq$ 36. Pruebe al nivel de significancia del 5%.

"Maldita sea", murmuró Laurel para sí. "Espero no haberme equivocado mucho con esa suposición. Si tengo suerte y los datos en realidad siguen una distribución normal, podría verificarlo. Por lo menos tengo que comunicárselo a Stan y Hal." Laurel no se emocionaba al tener que reconocer un error.

4. Verifique la suposición de que los datos siguen una distribución normal. Pruebe al nivel de significancia de 0.05.

5. Si la suposición de normalidad parece razonable, vuelva a estimar el número de vendedores que se necesitan para atender los teléfonos.

Stan miró los datos de Laurel con interés. "Sabía que algo no andaba bien con nuestras conclusiones originales, pero no podía decir qué era. Fue buena idea mantener a nuestros seis vendedores", sonrió.

Laurel suspiró aliviada.

"Sin embargo", continuó él, "esto trae a colación una nueva pregunta. Estos datos parecen indicar, y lo sé por experiencia, que definitivamente hay algunas horas pico. Me pregunto si podemos llegar a una solución menos costosa si utilizamos una combinación de representantes de ventas de tiempo completo y de tiempo parcial. ¿Tú qué crees?"

"Quizá tengas razón", asintió Laurel. "Déjame hacer unos cuantos cálculos más y le llevamos los resultados a Hal para que nos dé su opinión."

6. Calcule el promedio y la desviación estándar del número de llamadas recibidas durante cada hora. Suponiendo que, durante cada hora, el número de llamadas recibidas tienen una distribución normal, calcule los niveles recomendados de personal con el fin de estar 98% seguros de que un representante de ventas solamente tiene que atender ocho llamadas por hora. ¿Qué combinación de vendedores de tiempo completo y tiempo parcial parece ser apropiada?

Del libro de texto al mundo real

Control estadístico de procesos

Los profesionales de mercadotecnia utilizan la estadística para analizar datos y determinar la efectividad de diferentes técnicas de comercialización. Las agencias de investigación en mercadotecnia pueden recolectar datos comerciales mediante entrevistas intensivas por teléfono o en persona. Además, las encuestas por correo pueden proporcionar una opción de bajo costo relativo para la recolección de información en poblaciones ampliamente dispersas. Uno de los inconvenientes principales de las encuestas por correo es que su nivel de respuesta es, por lo general, más bajo que el de las entrevistas en persona o telefónicas; así, su sesgo por no recibir respuesta es más aguda.

En 1987, se realizó un experimento en Londres para ver si podían mejorar los niveles de respuesta a las encuestas por correo adjuntando un pequeño incentivo monetario y/o un folleto informativo junto con la encuesta. Antes de 1987, las agencias de investigación inglesas preferían realizar encuestas por teléfono, y únicamente el 4% de las compañías en Inglaterra habían utilizado incentivos monetarios en sus encuestas por correo. El experimento se diseñó para probar las siguientes hipótesis nulas:

H_1: La respuesta a las encuestas comerciales por correo es independiente de la inclusión de un incentivo monetario.
H_2: La respuesta a las encuestas comerciales por correo es independiente de la inclusión de un folleto informativo.

El experimento La muestra consistió en 159 altos ejecutivos de empresas de construcción. Los cuestionarios se asignaron aleatoriamente incluyendo con ellas:

1. incentivo monetario, sin folleto,
2. incentivo monetario, con folleto,
3. sin incentivo monetario, con folleto,
4. sin incentivo monetario, sin folleto.

En la carta que contenía una moneda de 20 centavos, utilizada como incentivo monetario, se sugería que el dinero podría utilizarse para comprar una taza de café y facilitar la tarea de llenar el cuestionario. La carta adjunta al folleto indicaba que contenía una explicación de la importancia de la encuesta para la investigación.

Tabla RW 11-1 Resultados ANOVA para la tasas de respuesta

Fuente de variación	Suma de cuadrados	GL	Media al cuadrado	*F*
Efectos principales	1.04	2	0.052	2.26
Incentivo monetario	0.87	1	0.87	3.76*
Folleto	0.16	1	0.16	0.71

*Denota nivel de significancia de 0.05.

Los resultados Se utilizó el análisis de varianza para probar los efectos de los tratamientos sobre las tasas de respuesta. Como verificación, se usó también la prueba ji-cuadrada para probar la asociación de los efectos principales sobre la tasa de respuesta. Esta tasa de respuesta global a la encuesta por correo fue 36.5%. Sin embargo, la tasa de respuesta asociada con la moneda de 20 centavos fue 44.2% (contra el 29.3% para las encuestas sin incentivo monetario) y con el folleto fue 33.3% (contra el 40.0% para las encuestas sin folleto). En la tabla MR11-1 se presentan los resultados de ANOVA. El aumento del 14.9% en la tasa de respuesta a las encuestas con moneda fue significativo para $\alpha = 0.05$. La diferencia folleto/no folleto no fue significativa. La prueba ji-cuadrada confirmó los resultados de ANOVA.

Relevancia práctica Los profesionales de Gran Bretaña han externado su escepticismo respecto a las encuestas por correo debido al sesgo por no recibir respuestas. El experimento indicó que un pequeño incentivo monetario podría mejorar las tasas de respuesta a las encuestas por correo sin afectar la calidad de la respuesta. Los folletos, que son costosos, no tienen un efecto significativo en la tasa de respuesta; de hecho, esta tasa fue mayor para los cuestionarios sin folleto. Esta evidencia sugiere que las encuestas por correo son un medio viable para obtener información de mercado y que el tiempo y el esfuerzo implicados en la producción de un folleto puede invertirse mejor en incentivos monetarios.

Fuente: David Jobber, Karl Birro y Stuart Sanderson, "A Factorial Investigation of Methods of Stimulating Response to a Mail Survey", *European Journal of Operational Research* 37 (1988): 158-163.

Repaso del capítulo

● Términos introducidos en el capítulo 11

Análisis de varianza (ANOVA) Técnica estadística utilizada para probar la igualdad de tres o más medias muestrales y hacer inferencias sobre si las muestras provienen de poblaciones que tienen la misma media.

Cociente *F* Cociente utilizado en el análisis de varianza, entre otras pruebas, para comparar la magnitud de dos estimaciones de la varianza de la población para determinar si las dos estimaciones son aproximadamente iguales; en ANOVA se usa el cociente de la varianza entre columnas dividido entre la varianza dentro de columnas.

Distribución *F* Familia de distribuciones diferenciadas por dos parámetros (gl del numerador, gl del denominador), utilizada principalmente para probar hipótesis sobre varianzas.

Distribución ji-cuadrada Familia de distribuciones de probabilidad, diferenciadas por sus grados de libertad, que se utiliza para probar varias hipótesis diferentes acerca de varianzas, proporciones y bondad de ajuste de distribuciones.

Frecuencias esperadas Frecuencias que esperaríamos ver en una tabla de contingencia o distribución de frecuencias si la hipótesis nula es verdadera.

Gran media La media para el grupo completo de sujetos provenientes de todas las muestras del experimento.

Prueba de bondad de ajuste Prueba estadística para determinar si existe una diferencia significativa entre una distribución de frecuencias observadas y una distribución de probabilidad teórica hipotética para describir la distribución observada.

Prueba de independencia Prueba estadística de proporciones de frecuencias que se usa para determinar si la pertenencia a las categorías de una variable es diferente como función de la pertenencia a las categorías de una segunda variable.

Tabla de contingencia Tabla que contiene R renglones y C columnas. Cada renglón corresponde a un nivel de una variable; cada columna, a un nivel de otra variable. Los elementos del cuerpo de la tabla son las frecuencias con que ocurre cada combinación de variables.

Varianza dentro de columnas Estimación de la varianza de la población basada en las varianzas dentro de k muestras, utilizando un promedio ponderado de k varianzas muestrales.

Varianza entre columnas Estimación de la varianza de la población derivada de la varianza entre las medias de las muestras.

● Ecuaciones introducidas en el capítulo 11

■ **11-1**
$$\chi^2 = \Sigma \frac{(f_o - f_e)^2}{f_e}$$

Esta fórmula dice que el *estadístico ji-cuadrada* (χ^2) es igual a la suma (Σ) que obtendremos si:

1. Restamos las frecuencias esperadas, f_e, de las frecuencias observadas, f_o, para cada categoría de la tabla de contingencia.
2. Elevamos al cuadrado cada diferencia.
3. Dividimos cada diferencia al cuadrado entre f_e.
4. Sumamos todos los resultados obtenidos en el paso 3.

■ **11-2** Número de grados de libertad = (número de renglones − 1)(número de columnas − 1)

Para calcular el número de *grados de libertad de una prueba de independencia ji-cuadrada*, multiplique el número de renglones (menos 1) por el número de columnas (menos 1).

■ **11-3**
$$f_e = \frac{RT \times CT}{n}$$

Con esta fórmula podemos calcular la frecuencia esperada para cualquier celda de una tabla de contingencia. *RT* es el total por renglón para el renglón que contiene a la celda, *CT* es el total por columna para la columna que contiene a la celda y n es el número total de observaciones.

■ **11-4**
$$s_{\bar{x}}^2 = \frac{\Sigma(\bar{x} - \bar{\bar{x}})^2}{k - 1}$$

Para calcular la *varianza entre las medias muestrales*, utilice esta fórmula.

■ **11-5**
$$\sigma^2 = \sigma_{\bar{x}}^2 \times n$$

La *varianza de la población* es igual al producto del cuadrado del error estándar de la media y el tamaño de la muestra.

■ **11-6**
$$\hat{\sigma}_b^2 = \frac{\Sigma n_j(\bar{x}_j - \bar{\bar{x}})^2}{k - 1}$$

Una estimación de la varianza de la población (la varianza entre columnas) puede obtenerse con esta ecuación. La ecuación se obtiene sustituyendo primero $s_{\bar{x}}^2$ por $\sigma_{\bar{x}}^2$ en la ecuación 11-5, y luego ponderando cada $(\bar{x}_j - \bar{\bar{x}})^2$ con su tamaño de muestra adecuado (n_j).

■ **11-7**
$$\hat{\sigma}_w^2 = \Sigma\left(\frac{n_j - 1}{n_T - k}\right)s_j^2$$

Una segunda estimación de la varianza de la población (la varianza dentro de las columnas) se obtiene con esta ecuación. Esta ecuación utiliza un promedio ponderado de todas las varianzas muestrales. En esta formulación, $n_T = \Sigma n_j$, el tamaño de muestra total.

11-8

$$F = \frac{\text{primera estimación de la varianza de la población basada en la varianza entre las medias muestrales}}{\text{segunda estimación de la varianza de la población basada en las varianzas dentro de las muestras}}$$

Este cociente permite comparar las dos estimaciones de la varianza de la población, calculadas con las ecuaciones 11-6 y 11-7. En una prueba de hipótesis basada en una distribución F, tenemos más probabilidad de aceptar la hipótesis nula si este cociente F o estadístico F tiene un valor cercano a 1. Conforme el cociente F aumenta, es más probable que se rechace la hipótesis nula.

11-9

$$F = \frac{\text{varianza entre columnas}}{\text{varianza dentro de columnas}} = \frac{\hat{\sigma}_b^2}{\hat{\sigma}_w^2}$$

Ésta es otra forma de plantear la ecuación 11-8, utilizando símbolos estadísticos para el numerador y el denominador del cociente F.

11-10 Número de grados de libertad en el numerador del cociente F = (número de muestras − 1)

Para realizar un análisis de varianza, calculamos el número de *grados de libertad en la varianza entre columnas* (el numerador del cociente F) restando 1 al número de muestras tomadas.

11-11 Número de grados de libertad en el denominador del cociente $F = \Sigma(n_j - 1) = n_T - k$

Esta ecuación sirve para calcular el número de grados de libertad en el denominador del cociente F. Éste es igual al tamaño total de las muestras, n_T, menos el número de muestras, k.

11-12

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

Con una varianza de población σ^2, el estadístico χ^2 dado por esta ecuación tiene una distribución ji-cuadrada con $n - 1$ grados de libertad. Este resultado es exacto si la población es normal, pero aun en muestras tomadas de poblaciones no normales, a menudo es una buena aproximación.

11-13

$$\sigma^2 = \frac{(n-1)s^2}{\chi^2}$$

Para obtener un intervalo de confianza para σ^2, se despeja σ^2 de la ecuación 11-12.

11-14

$$\sigma_L^2 = \frac{(n-1)s^2}{\chi_U^2} \leftarrow \text{Límite inferior de confianza}$$

$$\sigma_U^2 = \frac{(n-1)s^2}{\chi_L^2} \leftarrow \text{Límite superior de confianza}$$

Estas fórmulas dan los límites de confianza inferior y superior de un intervalo de confianza para σ^2. (Note que debido a que χ^2 aparece en el denominador, utilizamos χ_U^2 para encontrar σ_L^2, y χ_L^2 para encontrar σ_U^2.)

11-15

$$F = \frac{s_1^2}{s_2^2}$$

Este cociente tiene una distribución F con $n_1 - 1$ grados de libertad en el numerador y $n_2 - 1$ grados de libertad en el denominador. (Esto supone que las dos poblaciones están razonablemente bien descritas por distribuciones normales.) Se usa para probar hipótesis de dos varianzas de población.

11-16

$$F(n, d, \alpha) = \frac{1}{F(d, n, 1-\alpha)}$$

La tabla 6 del apéndice sólo da valores de F para pruebas de hipótesis de extremo superior, pero esta ecuación permite encontrar los valores F apropiados para pruebas de cola inferior y de dos colas.

Ejercicios de repaso

■ **11-57** La oficina de correos está preocupada por la variabilidad en el número de días que toma a una carta ir de la costa este a la costa oeste de Estados Unidos. Se envió por correo una muestra de cartas desde la costa este y se registró el tiempo que tardaron en llegar a su destino en la costa oeste. Los datos obtenidos son:

Tiempo para llegar al destino (en días)									
2.2	1.7	3.0	2.9	1.9	3.1	4.2	1.5	4.0	2.5

Encuentre un intervalo de confianza del 90% para la varianza en los tiempos de llegada.

■ **11-58** Para la siguiente tabla de contingencia, calcule las frecuencias observadas y esperadas y el estadístico ji-cuadrada. Establezca y pruebe las hipótesis apropiadas al nivel de significancia de 0.05.

	Actitud hacia la legislación social		
Ocupación	**A favor**	**Neutral**	**Opuesto**
Obrero	19	16	37
Oficinista	15	22	46
Profesional	24	11	32

■ **11-59** Los vendedores saben que los gustos difieren en las diferentes regiones del país. En el negocio de renta de automóviles, un experto de la industria opina que existen fuertes preferencias regionales en cuanto al tamaño del auto y da los siguientes datos para apoyar ese punto de vista:

	Región del país			
Tipo de auto preferido	**Noreste**	**Sureste**	**Noroeste**	**Suroeste**
Grande	105	120	105	70
Mediano	120	100	130	150
Todos los demás	25	30	15	30

a) Establezca las hipótesis nula y alternativa apropiadas.
b) ¿Apoyan los datos la opinión del experto para un nivel de significancia de 0.05?
c) ¿Y para un nivel de significancia de 0.20?

■ **11-60** ¿Qué distribución de probabilidad se utiliza en cada uno de los siguientes tipos de pruebas estadísticas?
a) Comparación de dos proporciones de población.
b) Valor de una sola varianza de población.
c) Comparación de tres o más medias de población.
d) Comparación de dos medias de población a partir de muestras dependientes pequeñas.

■ **11-61** Gap vende ropa informal para adulto con una estrategia de "valor": buena calidad a un precio razonable. En 1986, la compañía inició una nueva división, Gap Kids, con tiendas separadas que ofrecen ropa similar para niños de entre 2 y 12 años. Después de varios años de expansión rápida en Estados Unidos, Gap empezó a abrir tiendas en Canadá y el Reino Unido. Para el tercer trimestre de 1992, el número de tiendas en operación era:

	Estados Unidos	**Inglaterra**	**Canadá**
Gap (adultos)	822	20	31
Gap Kids	240	8	14

Fuente: Company data, informe de Salomon Bros., 23 de diciembre de 1992.

Pruebe para $\alpha = 0.01$ si Gap está poniendo el mismo énfasis al abrir tiendas Gap Kids en los tres países. ¿Por qué tiene sentido esta estrategia de negocios?

■ **11-62** Los inventores y las compañías internacionales saben del valor de una patente estadounidense para proteger sus ideas, y en años recientes, han recibido casi la mitad de las patentes otorgadas. A partir de los datos de la siguiente tabla, ¿ha habido un cambio significativo en la proporción de patentes originadas fuera de Estados Unidos en los últimos 10 años? Pruebe para $\alpha = 0.05$.

Patentes concedidas	**1981**	**1991**
Originadas en EUA	39,223	51,183
Originadas fuera de EUA	26,548	45,331

Fuente: U.S. Patent Office, informe de Business Week (18 de enero de 1993): 79.

■ **11-63** Existen 33 orquestas sinfónicas importantes en Estados Unidos. El número de conciertos dados por cada una durante 1989 se lista y resume en la siguiente tabla. No queda claro de manera inmediata si estas orquestas pueden considerarse representativas de una sola población o si existen varios tipos diferentes (pequeña, media y grande), diferenciadas por el número de conciertos que dan al año. Si existen diferentes tipos de orquestas, una compañía editora de música podría querer desarrollar diferentes programas de comercialización para manejarlas. Por ejemplo, las orquestas sinfónicas podrían tener representantes de ventas asignados a ellas, pero un solo representante podría encargarse de varias orquestas sinfónicas más pequeñas.

Para darse una primera idea de si las 33 orquestas se pueden considerar un solo grupo, utilice la información de la tabla para probar para $\alpha = 0.025$, si el número de conciertos dados en 1989 están bien descritos por una distribución normal con $\mu = 182.3$ (la media de la muestra) y $\sigma = 57$ (la desviación estándar de la muestra).

Número de conciertos (datos sin procesar)

325	300	267	263	250	230	216	215	200	200	200
200	190	185	185	180	180	180	180	175	175	164
160	160	157	150	135	120	115	105	100	84	70

Distribución de frecuencias

Clase	≤100	101-150	151-200	201-250	251-300	≥301
Frecuencia	3	5	17	4	3	1

Fuente: Richard Boyer y David Savageau, Places Rated Almanac *(Nueva York: Prentice Hall Travel, 1989): 226.*

■ **11-64** ¿Qué distribución de probabilidad se utiliza en cada una de las siguientes pruebas estadísticas?

a) Comparación de las medias de dos muestras pequeñas de poblaciones con varianza desconocida.
b) Comparación de dos varianzas de población.
c) Valor de una sola media de población basado en muestras grandes.
d) Comparación de tres o más proporciones de población.

■ **11-65** Las tiendas establecen sus precios, pero los fabricantes tienen interés en el precio de venta final como parte de su estrategia de promoción. El gerente de mercadotecnia de los bolígrafos marca C se queja de que el resultado de un recorte de precios excesivo en las tiendas es que la marca C se percibe como una "pluma corriente". El gerente de ventas contesta que "todos dan descuentos, todas las marcas, hasta cierto punto". Durante las llamadas de ventas recolectaron datos del precio final de cuatro marcas de bolígrafos, incluyendo la de ellos, en cinco tiendas diferentes. Para un nivel de confianza de 0.05, ¿existe una variación significativa en el precio entre las marcas?

	Precio (en centavos de dólar)		
Marca A	**Marca B**	**Marca C**	**Marca D**
61	52	47	67
55	58	52	63
57	54	49	68
60	55	49	59
62	58	57	65

■ **11-66** Una compañía publicitaria de espectaculares debe saber si hay volúmenes de tránsito significativamente distintos que pasen frente a anuncios situados en tres lugares en Newark, debido a que la compañía cobra precios diferentes dependiendo de la cantidad de autos que pasan frente al anuncio. Se mide el volumen del tránsito en los tres sitios durante intervalos de 5 minutos elegidos aleatoriamente. La tabla muestra los datos obtenidos. Al nivel de significancia de 0.05, ¿son iguales los volúmenes de autos que pasan frente los tres anuncios?

Volumen de tránsito

Anuncio 1	30	45	26	44	18	38	42	29	
Anuncio 2	29	38	36	21	36	18	17	30	32
Anuncio 3	32	44	40	43	24	28	18		

■ **11-67** Un inversionista está interesado en ver si existen diferencias significativas en las tasas de rendimiento de acciones, bonos y fondos mutuos. Ha tomado muestras aleatorias de cada tipo de inversión y ha registrado los siguientes datos:

	Tasa de rendimiento (porcentaje)					
Acciones	2.0	6.0	2.0	2.1	6.2	2.9
Bonos	4.0	3.1	2.2	5.3	5.9	
Fondos mutuos	3.5	3.1	2.9	6.0		

a) Establezca las hipótesis nula y alternativa.
b) Pruebe sus hipótesis al nivel de significancia de 0.05.
c) Establezca una conclusión explícita.

■ **11-68** Para la siguiente tabla de contingencia:
a) Construya una tabla de frecuencias observadas y esperadas.
b) Calcule el estadístico ji-cuadrada.
c) Establezca las hipótesis nula y alternativa.
d) Al nivel de significancia de 0.05, ¿deberá rechazarse la hipótesis nula?

Concurrencia a la iglesia	Nivel de ingresos Bajo	Mediano	Alto
Nunca	27	48	15
Ocasionalmente	25	63	14
Regularmente	22	74	12

■ **11-69** La Overseas Shipholding Group, Inc. (OSG), tiene tres tipos de embarcaciones: navíos cargueros, transportadores de productos petroleros (TPP) y buques-tanque. Algunos de estos barcos están contratados a largo plazo y transportan bienes para un proveedor durante varios años. Otros barcos se contratan por viaje. La ventaja principal de un contrato a largo plazo es que se tienen ingresos fijos, al costo de renunciar a la oportunidad de cobrar un precio más alto si las fuerzas del mercado ponen a las embarcaciones en una demanda alta. La existencia de contratos a largo plazo afectará las estimaciones de los analistas financieros respecto a los ingresos futuros de la OSG.

¿Los siguientes datos de frecuencias indican que los tres tipos de embarcaciones tienen la misma probabilidad de ser contratados a largo plazo? Pruebe para $\alpha = 0.10$.

	Navíos cargueros	TPP	Buques-tanque
Contrato a largo plazo	7	7	20
Sin contrato a largo plazo	15	10	4

Fuente: Overseas Shipholding Group, Inc., 1991. Informe anual.

11-70 Los promedios Dow-Jones para la industria, el transporte y los servicios de luz, agua y gas se basan en el precio de las acciones de 30 empresas industriales, 20 empresas de transporte y 15 de servicios de luz, agua y gas, de la Bolsa de Valores de Nueva York, que se consideran representativas de todas las compañías que forman parte de sus grupos. La tabla da una lista de los cambios en los precios de las acciones para estas 65 compañías el día 21 de junio de 1993. Para $\alpha = 0.05$, ¿es razonable concluir que los tres grupos tuvieron cambios promedio significativamente diferentes en los precios de sus acciones ese día?

	Promedio industrial		
Compañía	Cambio	Compañía	Cambio
Alcoa	+0.125	Goodyear	−0.125
Allied Signal	+1.625	IBM	−0.125
American Express	−0.375	International Paper	+0.125
AT&T	+0.250	McDonald's	−0.250
Bethlehem Steel	+0.500	Merck	0
Boeing	−0.375	Minnesota Mining & Mfg.	+1.375
Caterpillar	−1.500	JP Morgan	+0.375
Chevron	+1.000	Philip Morris	−0.125
Coca Cola	−0.250	Procter & Gamble	+0.375
Disney	0	Sears	+0.500

(Continúa)

Promedio industrial *(continuación)*

Compañía	Cambio	Compañía	Cambio
DuPont	−0.250	Texaco	+1.000
Eastmon Kodak	+0.250	Union Carbide	+0.125
Exxon	−0.125	United Technologies	+0.750
General Electric	+1.000	Westinghouse	0
General Motors	+1.125	Woolworth	+0.250

Promedio de transporte

Compañía	Cambio	Compañía	Cambio
AMR	−0.500	Federal Express	+0.375
Airborne Freight	0	Norfolk Southern	+0.250
Alaska Air	−0.125	Roadway Services	+0.250
American President	−0.250	Ryder System	+0.125
Burlington Northern	+0.625	Santa Fe Pacific	+1.000
CSX	+1.125	Southwest Airlines	−0.625
Carolina Freight	+0.125	UAL	−1.500
Consolidated Freightways	−0.125	Union Pacific	+0.375
Conrail	+0.625	USAir	0
Delta Air lines	−0.125	XTRA	0

Promedio de servicios de luz, agua y gas

Compañía	Cambio	Compañía	Cambio
American Electric Power	+0.375	Niagara Mohawk Power	−0.500
Arkla	+0.125	Pacific Gas & Electric	+0.125
Centerior Energy	−0.125	Panhandie Eastern	+0.625
Commonwealth Edison	+0.625	Peopies Energy	+0.375
Consolidated Edison	+0.250	Philadelphia Electric	+0.250
Consolidated Natural Gas	+0.250	Public Service	
Detroit Edison	+0.375	Enterprise Group	+0.250
Houston Inclustries	−0.125	SCEcorp	+0.125

Fuente: The Wall Street Journal *(22 de junio de 1993): C3.*

■ **11-71** Para la siguiente tabla de contingencia:

a) Construya una tabla de frecuencias observadas y esperadas.
b) Calcule el estadístico ji-cuadrada.
c) Establezca las hipótesis nula y alternativa.
d) Al nivel de significancia de 0.01, ¿deberá rechazarse la hipótesis nula?

	Grupo de edad			
Tipo de auto	**16-21**	**22-30**	**31-45**	**46+**
4 × 4	19	23	15	2
Auto deportivo	9	14	11	7
Compacto	6	8	7	9
Mediano	11	13	19	24
Grande	9	13	22	26

■ **11-72** Swami Zhami asegura tener poderes psíquicos. Dice que puede adivinar correctamente el palo (diamantes, picas, tréboles, corazones) de una carta escogida al azar con una probabilidad de 0.5. Debido a que las cartas se escogen aleatoriamente de una baraja, podemos suponer que las adivinanzas de Zhami son independientes. En 100 días escogidos al azar, Zhami hizo 10 adivinanzas, y se registró el número de adivinanzas correctas. Queremos ver si este número tienen una distribución binomial con $n = 10$ y $p = 0.5$. Se recolectaron los siguientes datos:

Número de adivinanzas correctas por día	0-2	3-5	6-8	9-10
Frecuencia del número de adivinanzas correctas	50	47	2	1

a) Establezca explícitamente las hipótesis nula y alternativa.

b) Pruebe sus hipótesis. Utilice $\alpha = 0.10$.
c) Si Zharni no tiene poderes psíquicos, entonces deberá tener una probabilidad de 0.25 de adivinar correctamente una carta. (¿Por qué?) Vea si el número de adivinanzas correctas tiene distribución binomial con $n = 10$ y $p = 0.25$.

■ **11-73** Existe cierta evidencia sociológica de que las mujeres como grupo son más variables que los hombres en sus actitudes y creencias. Una organización privada de investigación efectuó un sondeo de las actitudes de los hombres sobre ciertos aspectos y encontró que la desviación estándar de esta escala de actitud era 16 puntos. Un sociólogo aplicó la misma escala a un grupo de 30 mujeres y encontró que la varianza de la muestra era 400 puntos al cuadrado. Al nivel de significancia de 0.01, ¿existe razón para creer que en efecto las mujeres muestran una mayor variabilidad en esta escala de actitud?

■ **11-74** Jim Greek hace predicciones acerca del número de canastas que anotará su equipo de baloncesto favorito. Estamos interesados en probar si sus errores siguen una distribución normal con media 0 y varianza 16. Use los datos siguientes para establecer explícitamente las hipótesis nula y alternativa, pruébelas al nivel $\alpha = 0.05$.

Error	≤ -7	-6 a 0	1 a 6	≥ 7
Número de predicciones	5	45	45	5

■ **11-75** Los sicólogos a menudo se preguntan sobre los efectos de la tensión y la ansiedad en el desempeño en los exámenes. Se aplicó una prueba de actitud a dos grupos escogidos aleatoriamente de 18 estudiantes universitarios, un grupo se hallaba en situación sin tensión y en el otro de tensión. El experimentador espera que la tensión aumente la varianza de las calificaciones del examen, pues siente que algunos estudiantes tienen mejor desempeño bajo tensión, mientras que otros tienen reacciones adversas ante ella. Las varianzas calculadas para los dos grupos son $s_1^2 = 23.9$ para el grupo sin tensión y $s_2^2 = 81.2$ para el grupo bajo tensión. ¿Se confirma su hipótesis? Utilice el nivel de significancia de 0.05 para probar las hipótesis.

La tabla MR11-2 contiene información de 50 áreas metropolitanas de Estados Unidos con poblaciones entre 190,000 y 290,000 habitantes. Las variables de la tabla son:

ÁREA	El nombre del área.
ESTADO	La abreviatura postal del estado.
REGIÓN	Nueva Inglaterra (NE), Atlántico Medio (AM), Central Norte (CN), Atlántico Sur (AS), Central Sur (CS), Montaña (MN) o Pacífico (PA).
POB	La población (en miles).
SOLO	Número de casas con un solo habitante (en miles).
EDAD	Edad promedio de los residentes.
VENTAS	Ventas al menudeo en 1991 (en miles de dólares).
ICE	Ingreso medio de compra efectivo por casa (en dólares).
COMIDA	Venta total de alimentos y bebidas consumidas fuera de casa en 1991 (en miles de dólares).
AUTO	Ventas totales por distribuidor de automóviles en 1991 (en miles de dólares).

Utilice la información dada de SOLO, EDAD y VENTAS para responder los ejercicios 11-76 a 11-78. (El resto de la información se utilizará en los capítulos 12 y 13.)

11-76 Una de las estrategias de negocios con más éxito es la llamada de "traslado". Cuando un concepto ha sido probado exitosamente en un área, una compañía se mueve gradualmente a nuevas áreas geográficas. (Algunos ejemplos son los centros de cambio de combustible JiffyLube y las cafeterías Starbuck.) En la estrategia de traslado, las empresas intentan saturar de tiendas un área geográfica a la vez, en lugar de abrir primero en, digamos, las 12 áreas metropolitanas más grandes. Tener juntas las nuevas tiendas significa un ahorro en costos de supervisión y distribución.

Suponga que una compañía tiene un nuevo producto cuyo atractivo varía con los grupos de edad de los consumidores, y está decidiendo en qué región del país deberá empezar la estrategia de traslado. ¿Las medias de las edades promedio de las áreas metropolitanas significativamente son diferentes en las regiones Central Sur, Central Norte y Atlántico Sur? Pruebe al nivel $\alpha = 0.05$.

11-77 Continuando con las estrategias de traslado, suponga que el nuevo producto está dirigido a las casas con un solo residente. ¿Los números medios de estas casas por área metropolitana son significativamente diferentes (para $\alpha = 0.01$) en las siete regiones del país?

11-78 Para las 50 áreas metropolitanas listadas, ¿la distribución de ventas totales al menudeo está bien descrita por una distribución normal?

Tabla MR11-2

Datos seleccionados para 50 pequeñas áreas metropolitanas

Área	Estado	Región	Pob.	Solo	Edad	Ventas	ICE	Comida	Auto
Salem	OR	PA	286.4	25.6	34.2	1,928,316	28,947	216,166	319,221
Rockford	IL	NC	285.6	26.4	33.8	2,257,983	35,334	220,669	440,350
Evansville	IN	NC	280.2	28.6	34.5	2,110,287	29,197	224,795	179,264
Fayetteville	NC	SA	278.1	18.3	27.6	1,719,843	26,590	164,397	450,351
Erie	PA	MA	275.4	26.3	33.3	1,788,155	29,682	182,161	387,164
Lorain-Elyria	OH	NC	272.5	20.4	33.1	1,712,808	31,041	149,020	248,501
Provo-Orem	UT	MN	271.7	9.0	22.5	1,173,520	30,441	90,884	408,123
Fort Pierce	FL	SA	267.3	24.4	40.9	2,066,935	30,024	185,405	377,386
Brownsville-Harlingen	TX	SC	267.2	12.3	27.6	1,442,992	19,881	128,964	515,363
Reno	NV	MN	266.1	29.9	33.9	2,584,652	32,561	211,024	400,512
Poughkeepsie	NY	MA	265.1	20.7	33.8	2,261,077	44,063	174,207	304,383
Binghamton	NY	MA	264.1	26.1	34.3	1,905,840	31,276	169,295	344,267
Killeen-Temple	TX	SC	262.6	18.0	27.5	1,398,926	27,301	127,669	374,750
New London-Norwich	CT	NE	257.6	22.1	32.9	1,839,535	39,019	210,759	392,641
Vancouver	WA	PA	250.2	20.9	33.1	1,392,459	34,132	170,967	294,343
Charleston	WV	SA	249.9	26.4	36.7	1,825,939	26,982	172,694	390,643
South Bend-Mishawaka	IN	NC	249.5	25.1	33.3	1,988,189	30,211	205,020	266,560
Huntsville	AL	SC	248.2	23.3	32.3	1,938,007	35,828	185,320	376,236
Springfield	MO	NC	248.2	24.7	33.1	2,111,044	27,465	178,733	410,634
Savannah	GA	SA	247.5	23.2	32.3	1,912,718	27,537	197,581	313,408
Portland	ME	NE	247.4	24.7	34.1	3,058,969	34,729	301,664	474,168
Columbus	GA	SA	246.7	21.3	30.5	1,686,624	24,366	170,745	323,467
Tallahassee	FL	SA	241.8	24.3	29.7	1,850,854	27,203	196,719	348,439
Johnstown	PA	MA	238.2	23.7	37.6	1,315,957	23,617	112,797	296,473
Duluth	MN	NC	237.1	27.5	36.1	1,786,197	25,546	154,514	543,951
Santa Cruz	CA	PA	234.5	20.7	33.2	1,899,989	38,972	234,696	259,757
Anchorage	AK	PA	232.4	19.8	29.8	2,518,807	46,573	419,996	390,938
Boulder-Longmont	CO	MN	230.0	24.2	31.7	2,012,188	37,543	222,162	365,778
Lubbock	TX	SC	226.5	21.5	29.3	2,025,939	26,553	204,994	357,671
Kalamazoo	MI	NC	226.1	21.3	31.4	1,799,240	33,068	184,597	487,152
Hickory-Morganton	NC	SA	226.0	19.6	35.0	1,563,932	28,795	142,331	413,278
Roanoke	VA	SA	225.0	24.2	36.8	2,090,618	29,611	179,906	266,968
Niagara Falls	NY	MA	220.7	22.6	35.0	1,462,303	30,540	137,316	457,468
Bradenton	FL	SA	220.5	26.1	43.5	1,671,344	26,626	190,590	430,663
Galveston-Texas City	TX	SC	219.8	20.4	32.8	1,587,897	30,652	171,136	373,363
Lincoln	NE	NC	217.6	23.6	31.1	1,579,558	33,698	172,109	355,544
Boise	ID	MN	210.9	19.0	32.1	1,550,437	33,548	172,069	298,075
Lafayette	LA	SC	209.9	18.0	29.8	1,667,771	26,508	176,663	288,170
Gainesville	FL	SA	208.9	22.5	29.4	1,658,143	23,460	168,526	282,377
Ocala	FL	SA	206.4	19.4	40.4	1,482,843	23,116	144,218	305,251
Bremerton	WA	PA	199.4	16.4	32.1	1,178,139	34,372	117,136	297,804
Biloxi-Gulfport	MS	SC	198.5	17.6	31.6	1,180,632	24,294	135,468	551,037
Green Bay	WI	NC	197.4	18.0	31.8	1,758,825	34,218	161,551	263,043
Fort Collins-Loveland	CO	MN	196.4	17.4	31.2	1,430,687	30,815	190,590	204,079
St. Cloud	MN	NC	196.0	13.9	28.7	2,613,718	30,257	118,871	477,988
Brazoria	TX	SC	194.3	12.3	31.1	1,135,072	36,422	90,314	168,512
Yakima	WA	PA	191.7	15.3	31.8	1,233,242	25,603	132,055	246,556
Springfield	IL	NC	191.4	22.8	34.7	1,513,666	34,403	134,863	223,719
Waco	TX	SC	191.2	18.6	31.5	1,430,358	25,958	137,323	366,036
Chico	CA	PA	190.1	19.3	34.2	1,139,717	24,138	125,090	441,586

Fuente: Adoptado de Sales & Marketing Management *(24 de agosto de 1992).*

■ **11-79** Con el fin de determinar la respuesta de las mujeres profesionales a las marcas de guardarropa de trabajo, On the Job, una boutique local, encuestó grupos de mujeres agentes de bienes raíces, secretarias, empresarias y ejecutivas de cuenta con respecto al estilo de ropa que usaban con más frecuencia (A, B, C, D). Se recolectaron los siguientes datos:

	Estilo			
Ocupación	**A**	**B**	**C**	**D**
Agente de bienes raíces	5	7	6	8
Secretaria	10	15	12	8
Empresaria	8	12	21	25
Ejecutiva de cuenta	12	14	20	25

Al nivel de significancia de 0.10, pruebe si el estilo que prefieren las mujeres depende de sus actividades.

■ **11-80** En el desarrollo de nuevas medicinas para el tratamiento de la ansiedad, es importante verificar el efecto de los medicamentos sobre varias funciones motrices, una de las cuales es conducir un automóvil. La compañía farmacéutica Confab está probando cuatro ansiolíticos diferentes respecto a su efecto sobre las capacidades para conducir. Los sujetos toman una prueba de manejo simulada y su calificación refleja los errores cometidos. Los errores más graves producen calificaciones más altas. Los resultados de estas pruebas se presentan en la siguiente tabla:

Medicina 1	245	258	239	241	
Medicina 2	277	276	263	274	
Medicina 3	215	232	225	247	226
Medicina 4	241	253	237	246	240

Al nivel de significancia de 0.05, ¿los cuatro medicamentos afectan de manera diferente las habilidades de manejo?

■ **11-81** James Clark acaba de adquirir dos fábricas de papel y está preocupado porque tienen una variabilidad significativamente diferente en sus producciones, aun cuando las dos plantas producen aproximadamente la misma cantidad promedio de papel cada día. La siguiente información se obtuvo para ver si las preocupaciones del señor Clark son justificadas. Al nivel de significancia $\alpha = 0.02$, ¿las dos plantas revelan la misma varianza en su producción?

Planta	n	s^2
Número 1	31	984 toneladas al cuadrado
Número 2	41	1,136 toneladas al cuadrado

■ **11-82** Los costos de combustible son importantes para alcanzar la rentabilidad en las líneas aéreas. Un pequeño transportista regional opera tres tipos de equipo y recolectó los siguientes datos de sus 14 aviones, expresados en costo de combustible (en centavos de dólar) por asiento-milla disponible.

Tipo A	7.3	8.3	7.6	6.8	8.0	
Tipo B	5.6	7.6	7.2			
Tipo C	7.9	9.5	8.7	8.3	9.4	8.4

A un nivel de significancia de 0.01, ¿se puede concluir que no existe una diferencia verdadera en los costos del combustible entre los tipos de aviones?

■ **11-83** Una muestra aleatoria de 50 jugadores de las ligas mayores de béisbol dio los siguientes promedios de bateo:

	Liga Americana				**Liga Nacional**		
Jugador	**Equipo**	**División**	**Promedio**	**Jugador**	**Equipo**	**División**	**Promedio**
Martínez	Seattle	Oeste	0.345	Grudsielanek	Montreal	Este	0.346
Baines	Chicago	Centro	0.325	Mabry	St. Louis	Centro	0.330
Hamilton	Texas	Oeste	0.320	Lansing	Montreal	Este	0.324
Córdova	Minnesota	Centro	0.311	Joyner	San Diego	Oeste	0.322
Thorne	Cleveland	Centro	0.309	Rodríguez	Montreal	Este	0.311
Buhner	Seattle	Oeste	0.305	Kendall	Pittsburgh	Centro	0.303
Franco	Cleveland	Centro	0.304	McGee	St. Louis	Centro	0.302
González	Texas	Oeste	0.294	Hunter	Houston	Centro	0.295

(*continúa*)

Liga Americana				Liga Nacional			
Jugador	**Equipo**	**División**	**Promedio**	**Jugador**	**Equipo**	**División**	**Promedio**
Roberts	Kansas	Centro	0.293	Battle	Oakland	Oeste	0.207
Suhoff	Baltimore	Este	0.293	Hamelin	Kansas	Centro	0.196
Cordero	Boston	Este	0.287	Castilla	Colorado	Oeste	0.277
Tettleton	Texas	Oete	o.287	Zeile	Philadelphia	Este	0.276
Leunitz	Nueva York	Este	0.278	Finley	San Diego	Oeste	0.267
Canseco	Boston	Este	0.275	McRae	Chicago	Centro	0.265
Ripken	Baltimore	Este	0.269	Butler	Los Ángeles	Oeste	0.265
Nieves	Detroit	Este	0.269	Thompson	San Francisco	Oeste	0.254
Nixon	Toronto	Este	0.266	Brogna	Nueva York	Este	0.252
Baerga	Cleveland	Centro	0.256	Branson	Cincinnati	Centro	0.243
Snow	California	Oeste	0.254	Bell	Pittsburgh	Centro	0.242
Davis	Seattle	Oeste	0.245	Veras	Florida	Este	0.241
Hoiles	Baltimore	Este	0.237	Santiago	Philadelphia	Este	0.228
Howard	Kansas	Centro	0.236	Johnson	Florida	Este	0.226
Young	Oakland	Oeste	0.225	Cedeno	San Diego	Oeste	0.224
Becker	Minnesota	Centro	0.208	Fonville	Los Ángeles	Oeste	0.216

Fuente: The News & Observer *(2 de junio de 1996): C9.*

¿Existe una evidencia significativa para concluir, a un nivel de significancia del 5%, que los jugadores en las seis divisiones tienen diferentes medias de promedio de bateo?

■ **11-84** Dick y Dave discutían sobre la variabilidad en las ligas mayores. Dick insistía en que no había diferencia en esa variabilidad entre las dos ligas. Dave aseguraba de la misma manera que había más variabilidad en la Liga Americana. Use los datos del ejercicio 11-83 (con $\alpha = 0.10$) para resolver la disputa.

■ **11-85** Un domingo, las secciones de clasificados en el *News & Observer* (N&O) de Raleigh, *The Chapel Hill News* (CHN) y el *Village Advocate* (VA) de Chapel Hill contenían los siguientes números de anuncios de venta de autos nacionales, autos extranjeros y camionetas/camiones ligeros.

	N&O	**CHN**	**VA**
Autos nacionales	543	32	36
Autos extranjeros	576	59	73
Camionetas/camiones ligeros	494	20	31

Fuente: The Newa & Observer *(2 de junio de 1996): I42-54*; The Chapel Hill News *(2 de junio de 1996): B9*; Village Advocate *(Chapel Hill, NC) (2 de junio de 1996): 4, 8, 17, 26, 27.*

a) Para $\alpha = 0.01$, pruebe si las proporciones de los tres tipos de anuncios varían significativamente entre los tres periódicos.

b) ¿Ayuda su conclusión del inciso a) para decidir qué periódico consultar si vive en Chapel Hill y está interesado en comprar un auto usado? Explique su respuesta.

capítulo 12

REGRESIÓN SIMPLE Y CORRELACIÓN

Objetivos

- Conocer cuántas decisiones comerciales dependen del conocimiento de la relación específica entre dos o más variables
- Utilizar diagramas de dispersión para visualizar la relación entre dos variables
- Emplear el análisis de regresión para estimar la relación entre dos variables
- Utilizar la ecuación de estimación de mínimos cuadrados para predecir valores futuros de la variable dependiente
- Aprender cómo el análisis de correlación describe el grado en el cual dos variables están relacionadas linealmente entre sí
- Comprender el coeficiente de determinación como una medida de la fuerza de la relación entre dos variables
- Conocer las limitaciones de la regresión y del análisis de correlación y las advertencias sobre su uso

Contenido del capítulo

12.1 Introducción 510
12.2 Estimación mediante la recta de regresión 516
12.3 Análisis de correlación 535
12.4 Inferencias sobre parámetros de población 545
12.5 Uso del análisis de regresión y correlación: limitaciones, errores y advertencias 551

- Estadística en el trabajo 553
- Ejercicio de base de datos computacional 553
- Del libro de texto al mundo real 554
- Términos introducidos en el capítulo 12 555
- Ecuaciones introducidas en el capítulo 12 555
- Ejercicios de repaso 557

El vicepresidente de investigación y desarrollo (ID) de una gran compañía química y de fabricación de fibras cree que las ganancias anuales de la empresa dependen de la cantidad gastada en ID. El nuevo presidente de la compañía no está de acuerdo y ha solicitado pruebas. Los datos de seis años son los siguientes:

Año	Millones gastados en investigación y desarrollo	Ganancia anual (millones)
1990	2	20
1991	3	25
1992	5	34
1993	4	30
1994	11	40
1995	5	31

El vicepresidente de ID desea una ecuación para pronosticar los beneficios anuales derivados de la cantidad presupuestada para ID. Con los métodos de éste capítulo, podremos proporcionarle esa herramienta para la toma de decisiones y orientarlo respecto a la precisión que puede esperar al usarla. ■

12.1 Introducción

Relación entre variables

Todos los días, los administradores toman decisiones personales y profesionales basadas en predicciones de sucesos futuros. Para hacer estos pronósticos, se basan en la relación (intuitiva y calculada) entre lo que ya se sabe y lo que se debe estimar. Si los responsables de la toma de decisiones pueden determinar cómo lo conocido se relaciona con el evento futuro, pueden ayudar considerablemente al proceso de toma de decisiones. Ése es el objetivo de este capítulo: cómo determinar la *relación entre variables*.

Diferencia entre la ji-cuadrada y los temas de este capítulo

En el capítulo 11, utilizamos pruebas de ji-cuadrada de independencia para determinar si existía una relación estadística entre dos variables. La prueba ji-cuadrada nos dice *si* existe tal relación, pero no nos dice *cuál* es esa relación. **Los análisis de regresión y correlación nos mostrarán cómo determinar tanto la naturaleza como la fuerza de una relación entre dos variables.** De esta forma, aprenderemos a pronosticar, con cierta precisión, el valor de una variable desconocida basándonos en observaciones anteriores de ésa y otras variables.

Origen de los términos regresión y regresión múltiple

El término *regresión* fue utilizado por primera vez como un concepto estadístico en 1877 por sir Francis Galton, quien llevó a cabo un estudio que mostró que la estatura de los niños nacidos de padres altos tiende a retroceder o "regresar" hacia la estatura media de la población. Designó la palabra *regresión* como el nombre del proceso general de predecir una variable (la estatura de los niños) a partir de otra (la estatura del padre o de la madre). Más tarde, los estadísticos acuñaron el término *regresión múltiple* para describir el proceso mediante el cual se utilizan varias variables para predecir otra.

Desarrollo de una ecuación de estimación

En el *análisis de regresión*, desarrollaremos una *ecuación de estimación*, esto es, una fórmula matemática que relaciona las variables conocidas con la variable desconocida. Después de conocer el patrón de esta relación, podremos aplicar el *análisis de correlación* para determinar el grado en el que las variables se relacionan. El análisis de correlación, entonces, nos indica qué tan bien la ecuación de estimación describe realmente la relación.

Tipos de relaciones

Variables independientes y dependientes

Los análisis de regresión y de correlación se basan en la relación, o asociación, entre dos (o más) variables. La variable (o variables) conocida(s) se llaman *variable(s) independiente(s)*; la que tratamos de predecir es la *variable dependiente.*

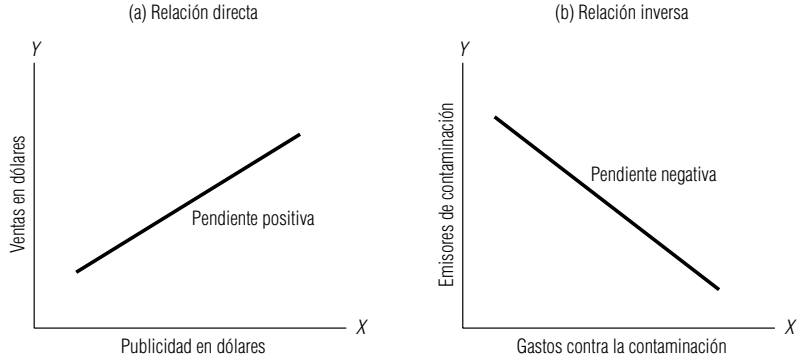


FIGURA 12-1

Relaciones directas e inversas entre la variable independiente *X* y la variable dependiente *Y*

Los científicos saben, por ejemplo, que existe una relación entre las ventas anuales de latas de aerosoles y la cantidad de fluorocarburos liberados a la atmósfera cada año. Si estudiáramos esta relación, "el número de latas de aerosol vendidas cada año" sería la variable independiente y "la cantidad de fluorocarburos liberados anualmente" sería la variable dependiente.

Consideremos otro ejemplo. Los economistas pueden basar sus predicciones del producto interno bruto anual, o PIB, en el gasto final de consumo dentro de la economía. Por tanto, "el consumo final" es la variable independiente y "el PNB" la variable dependiente.

En regresión, podemos tener sólo una variable dependiente en la ecuación de estimación. Sin embargo, podemos usar más de una variable independiente. A menudo, cuando agregamos variables independientes, mejoramos la exactitud de nuestra predicción. Los economistas, por ejemplo, con frecuencia añaden una segunda variable independiente, "el nivel de gasto de inversión", para mejorar su estimación del PIB.

Relación directa entre *X* y *Y*

Los dos ejemplos de fluorocarburos y PIB son ilustraciones de asociaciones directas entre variables independientes y dependientes. Al incrementarse la variable independiente, la variable dependiente también lo hace. De manera similar, esperamos que las ventas de una compañía se incrementen al aumentar el presupuesto de publicidad. Podemos graficar una *relación directa* de este tipo colocando la variable independiente en el eje *X* y la variable dependiente en el eje *Y*. La gráfica (a) de la figura 12-1 muestra esto. Note cómo la pendiente de la recta sube cuando *X* toma valores cada vez más grandes. Se dice que la pendiente de esta recta es *positiva*, porque *Y* crece si *X* crece.

Relación inversa entre *X* y *Y*

Las relaciones pueden ser *inversas* en vez de directas. En estos casos, la variable dependiente disminuye al aumentar la variable independiente. El gobierno supone que existe una asociación inversa entre un mayor gasto anual de una compañía en dispositivos anticontaminantes y menores emisiones contaminantes. La gráfica (b) de la figura 12-1 ilustra este tipo de relación, que se caracteriza por una pendiente *negativa* (la variable dependiente *Y* disminuye al aumentar la variable independiente *X*).

A menudo encontramos una *relación causal entre variables, esto es, la variable independiente "causa" cambios en la variable dependiente.* Éste es el caso en el ejemplo de la contaminación. Pero en muchos casos, otros factores ocasionan los cambios tanto en las variables dependientes como en las independientes. Podríamos predecir las ventas de aretes de diamantes observando las de Cadillacs nuevos, pero no podríamos decir que una origina a la otra. Más bien, nos damos cuenta que otro factor, como el nivel de ingresos disponibles, es la causa de los niveles de ventas tanto de Cadillacs como de aretes de diamantes.

Relaciones de asociación, no de causa y efecto

Por esta razón, es importante considerar que las relaciones encontradas por la regresión son relaciones de asociación, pero no necesariamente de causa y efecto. A menos que tenga razones específicas para creer que los valores de la variable dependiente se originan por los valores de las variables independientes, no infiera causalidad en las relaciones encontradas por la regresión.

Diagramas de dispersión

Diagrama de dispersión

El primer paso para determinar si existe una relación entre dos variables es examinar la gráfica de los datos observados (o conocidos). Esta gráfica, o dibujo, se llama *diagrama de dispersión*.

Tabla 12-1

Calificaciones de estudiantes en exámenes de admisión y promedios de generales acumulados al graduarse

Estudiante	A	B	C	D	E	F	G	H
Calificaciones de examen de admisión (100 = máxima calificación posible)	74	69	85	63	82	60	79	91
Promedio general acumulado (4.0 = A)	2.6	2.2	3.4	2.3	3.1	2.1	3.2	3.8

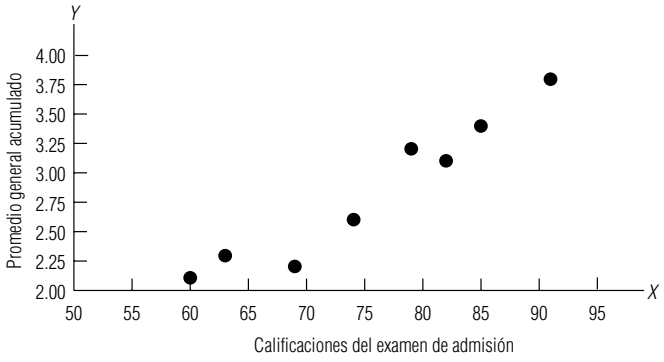


FIGURA 12-2

Diagrama de dispersión de las calificaciones de estudiantes en exámenes de admisión graficadas contra el promedio general acumulado

Un diagrama de dispersión nos puede dar dos tipos de información. Visualmente, podemos identificar patrones que indiquen que las variables están relacionadas. Si esto sucede, podemos ver qué tipo de línea, o ecuación de estimación, describe esta relación.

Desarrollaremos y utilizaremos un diagrama de dispersión específico. Suponga que el director de admisiones de una universidad nos pide determinar si existe una relación entre las calificaciones de un estudiante en su examen de admisión y su promedio general al graduarse. El director ha reunido una muestra aleatoria de datos de los registros de la universidad. La tabla 12-1 contiene esta información.

Transferencia de información tabular a una gráfica

Para comenzar, debemos transferir la información de la tabla 12-1 a una gráfica. Puesto que el director desea utilizar las calificaciones de los exámenes para pronosticar éxitos en la universidad, hemos colocado el promedio de calificaciones acumulado (la variable dependiente) en el eje vertical o *Y*, y la calificación del examen de admisión (la variable independiente) en el eje horizontal o *X*. La figura 12-2 nos muestra el diagrama de dispersión completo.

Trazo, o "ajuste", de una línea recta a través del diagrama de dispersión

A primera vista se sabe por qué llamamos así al diagrama de dispersión. El patrón de puntos resulta al registrar cada par de datos de la tabla 12-1 como un punto. Cuando vemos todos estos puntos juntos, podemos visualizar la relación que existe entre las dos variables. Como resultado, podemos trazar, o "ajustar" una línea recta a través de nuestro diagrama de dispersión para representar la relación; la figura 12-3 ilustra esto. Es común intentar trazar estas líneas de forma tal que un número igual de puntos caiga en cada lado de la línea.

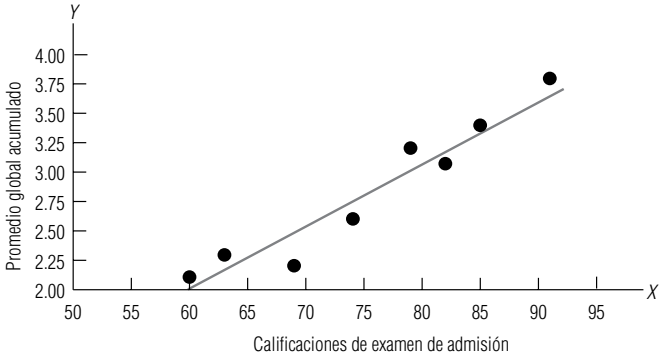


FIGURA 12-3

Diagrama de dispersión en donde la línea recta representa la relación entre *X* y *Y* "ajustada"

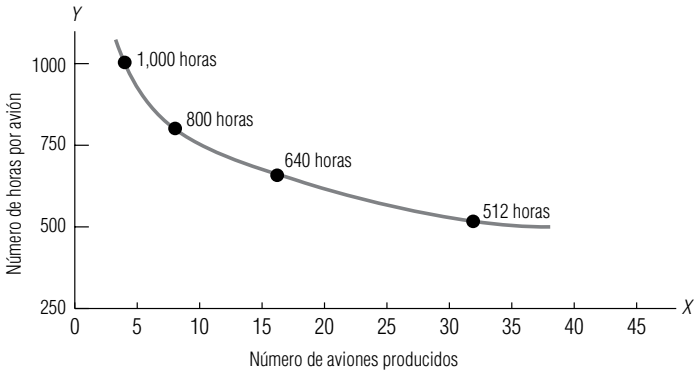


FIGURA 12-4

Relación curvilínea entre el tiempo de construcción de una nuevo avión y el número de unidades producidas

Interpretación de la línea recta

En este caso, la línea trazada a través de los puntos representa una relación directa, porque *Y* se incrementa al aumentar *X*. Como los puntos están relativamente cerca de esta línea, podemos decir que existe un alto grado de asociación entre las calificaciones de exámenes y el promedio de calificaciones acumulativo. En la figura 12-3, podemos ver que la relación descrita por los puntos está bien descrita por una línea recta. Por tanto, podemos decir que es una relación *lineal*.

Relaciones curvilíneas

La relación entre las variables *X* y *Y* también puede tomar la forma de una curva. Los especialistas en estadística la llaman relación *curvilínea*. Los empleados de muchas industrias, por ejemplo, experimentan lo que se denomina "curva de aprendizaje", es decir, al fabricar un nuevo producto, el tiempo requerido para producir una unidad se reduce en alguna proporción fija al duplicarse el número total de unidades. Una industria de este tipo es la aviación. El tiempo de fabricación por unidad de una nueva aeronave tiende a disminuir un 20% cada vez que se duplica el número de nuevos aviones terminados. La figura 12-4 ilustra la relación curvilínea de este fenómeno de "curva de aprendizaje".

La dirección de la curva puede indicar si la relación curvilínea es directa o inversa. La curva de la figura 12-4 describe una relación inversa porque *Y* disminuye al aumentar *X*.

Repaso de las relaciones posibles

Para repasar las relaciones posibles en un diagrama de dispersión, examinemos las gráficas de la figura 12-5. Las gráficas (a) y (b) muestran relaciones lineales directas e inversas. Las gráficas (c) y (d) son ejemplos de relaciones curvilíneas que indican asociaciones directas e inversas entre variables, respectivamente. La gráfica (e) ilustra una relación lineal inversa con un patrón de puntos ampliamente disperso. Esta mayor dispersión indica que existe menor grado de asociación entre las va-

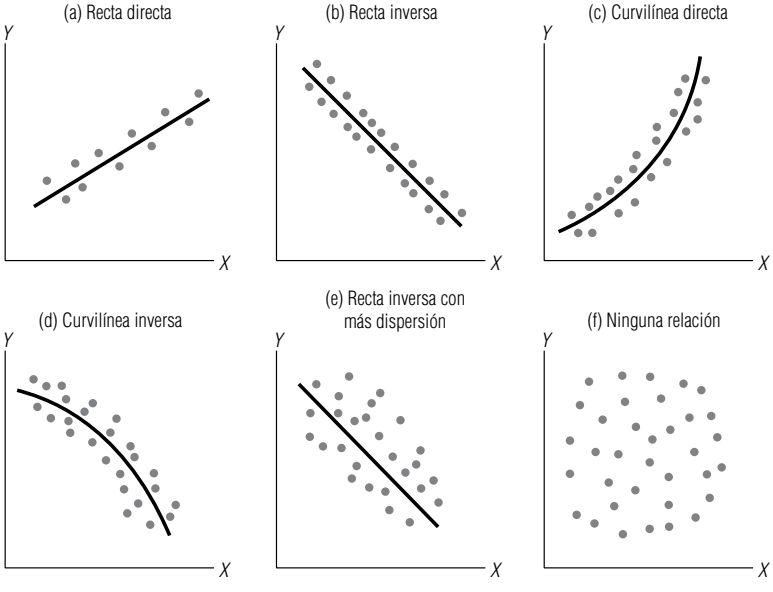


FIGURA 12-5

Relaciones posibles entre *X* y *Y* en diagramas de dispersión

riables independiente y dependiente que el existente en la gráfica (b). El patrón de puntos en la gráfica (*f*) parece indicar que no existe relación entre las dos variables; por tanto, conocer el pasado referente a una variable no nos permitirá pronosticar ocurrencias futuras de la otra.

Ejercicios 12.1

Ejercicios de autoevaluación

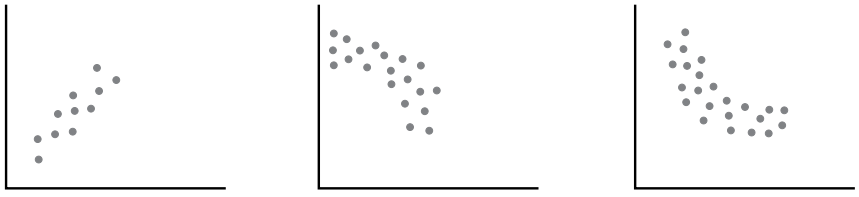
EA 12-1 Un instructor está interesado en saber cómo se relaciona el número de estudiantes ausentes con la temperatura media del día. Usó una muestra aleatoria de 10 días para el estudio. Los siguientes datos indican el número de estudiantes ausentes (AUS) y la temperatura media (TEMP) para cada día.

AUS	8	7	5	4	2	3	5	6	8	9
TEMP	10	20	25	30	40	45	50	55	59	60

a) Establezca la variable dependiente (*Y*) y la variable independiente (*X*).
b) Dibuje un diagrama de dispersión para estos datos.
c) ¿La relación entre las variables parece lineal o curvilínea?
d) ¿Qué tipo de curva puede dibujar a través de los datos?
e) ¿Cuál es la explicación lógica para la relación observada?

Conceptos básicos

- **12-1** ¿Qué es el análisis de regresión?
- **12-2** En el análisis de regresión, ¿qué es una ecuación de estimación?
- **12-3** ¿Cuál es el propósito del análisis de correlación?
- **12-4** Defina qué son las relaciones directas e inversas.
- **12-5** ¿A qué se refiere el término *relación causal*?
- **12-6** Explique la diferencia entre relaciones lineales y curvilíneas.
- **12-7** Explique por qué y cómo se construye un diagrama de dispersión.
- **12-8** ¿Qué es análisis de regresión múltiple?
- **12-9** Para cada uno de los siguientes diagramas de dispersión, indique si existe una relación y, en caso afirmativo, si es de tipo directo o inverso, y si es lineal o curvilínea.



Aplicaciones

- **12-10** Un profesor intenta mostrar a sus estudiantes la importancia de los exámenes cortos, aun cuando el 90% de la calificación final esté determinada por los exámenes parciales. Él cree que cuanto más altas sean las calificaciones de los exámenes cortos, más alta será la calificación final. Seleccionó una muestra aleatoria de 15 estudiantes de su clase con los siguientes datos:

Promedio de exámenes cortos	Promedio final
59	65
92	84
72	77
90	80
95	77

(Continúa)

Promedio de exámenes cortos	Promedio final
87	81
89	80
77	84
76	80
65	69
97	83
42	40
94	78
62	65
91	90

a) Establezca la variable dependiente (Y) y la variable independiente (X).
b) Dibuje un diagrama de dispersión para estos datos.
c) ¿La relación entre las variables parece lineal o curvilínea?
d) ¿Parece justificarse la idea del profesor? Explique su razonamiento.

■ **12-11** William Hawkins, vicepresidente de personal de la International Motors, trabaja en la relación entre el salario de un trabajador y el porcentaje de ausentismo. Hawkins dividió el intervalo de salarios de International en 12 grados o niveles (1 es el de menor grado, 12 el más alto) y después muestreó aleatoriamente a un grupo de trabajadores. Determinó el grado de salario de cada trabajador y el número de días que ese empleado había faltado en los últimos 3 años.

Categoría de salario	11	10	8	5	9	9	7	3
Ausencias	18	17	29	36	11	26	28	35
Categoría de salario	11	8	7	2	9	8	6	3
Ausencias	14	20	32	39	16	26	31	40

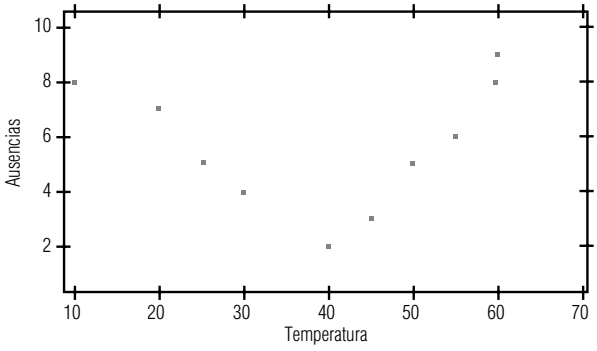
Elabore un diagrama de dispersión para estos datos e indique el tipo de relación.

■ **12-12** El Instituto Nacional de Ciencias para la Salud Ambiental (NIEHS, por sus siglas en inglés) ha estudiado las relaciones estadísticas entre muchas variables diferentes y el resfriado común. Una de las variables analizadas es el uso de pañuelos desechables (X) y el número de días de síntomas de resfrío mostrados (Y) por siete personas en un periodo de 12 meses. ¿Qué relación, si la hay, parece existir entre las dos variables? ¿Indica esto algún efecto causal?

X	2,000	1,500	500	750	600	900	1,000
Y	60	40	10	15	5	25	30

Soluciones a los ejercicios de autoevaluación

EA 12-1 a) Se desea ver si las ausencias (AUS) dependen de la temperatura (TEMP).
b)



c) Curvilínea.
d) Una curva cuadrática (parábola).
e) Cuando hace mucho frío o mucho calor hay muchos ausentes. Para temperaturas moderadas, no hay tantos estudiantes ausentes.

12.2 Estimación mediante la recta de regresión

Cálculo de la línea de regresión usando una ecuación

En los diagramas de dispersión que hemos utilizado hasta ahora, se colocaron las *líneas de regresión* ajustando las líneas visualmente entre los puntos de datos. En esta sección, aprenderemos a calcular la línea de regresión de manera más precisa, usando una ecuación que relaciona las dos variables matemáticamente. Aquí, examinaremos sólo relaciones lineales entre dos variables; estudiaremos las relaciones entre más de dos variables en el siguiente capítulo.

Ecuación para una línea recta

La ecuación para una línea recta donde la variable dependiente Y está determinada por la variable independiente X es:

Ecuación para una línea recta

$$Y = a + bX \qquad [12\text{-}1]$$

Variable dependiente: Y — Variable independiente: X — Variable ordenada: a — Pendiente de la recta: b

Interpretación de la ecuación

Usando esta ecuación, podemos tomar un valor dado de X y calcular el valor de Y. La a se denomina la "ordenada Y" porque su valor es el punto en el cual la línea de regresión cruza el eje Y, es decir, el eje vertical. La b en la ecuación 12-1 es la "pendiente" de la recta. Representan qué tanto cada cambio de una unidad de la variable independiente X hace que cambie la variable dependiente Y. Tanto a como b son *constantes* numéricas porque para cualquier línea recta dada, sus valores no cambian.

Cálculo de Y a partir de X usando la ecuación de la recta

Supongamos que sabemos que a es 3 y b es 2. Determinemos cuál sería Y para X igual a 5. Al sustituir los valores de a, b y X en la ecuación 12-1, encontramos que el valor correspondiente de Y es

$$\begin{aligned} Y &= a + bX \qquad [12\text{-}1] \\ &= 3 + 2(5) \\ &= 3 + 10 \\ &= 13 \leftarrow \text{Valor de } Y \text{ dada } X = 5 \end{aligned}$$

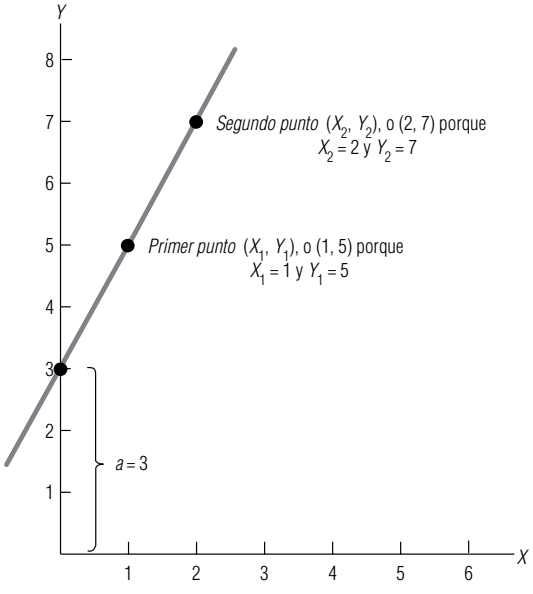


FIGURA 12-6

Línea recta con pendiente positiva, con la ordenada Y y dos puntos en la línea designada

Uso de la ecuación de estimación para una línea recta

Búsqueda de los valores de *a* y *b*

¿Cómo podemos encontrar los valores de las constantes numéricas, *a* y *b*? Para ilustrar este proceso, se usará la recta de la figura 12-6.

Podemos encontrar *a* visualmente (la ordenada *Y*) localizando el punto donde la recta cruza el eje *Y*. En la figura 12-6, esto sucede cuando $a = 3$.

Para encontrar la pendiente de la recta, *b*, debemos determinar cómo cambia la variable dependiente, *Y*, al cambiar la variable independiente, *X*. Podemos empezar por elegir dos puntos sobre la línea de la figura 12-6. Ahora, debemos encontrar los valores de *X* y *Y* (las *coordenadas*) de ambos puntos. Podemos llamar a las coordenadas de nuestro primer punto (X_1, Y_1) y (X_2, Y_2) a las del segundo. Al examinar la figura 12-6, podemos ver que (X_1, Y_1) = (1, 5) y (X_2, Y_2) = (2, 7). Entonces podemos calcular el valor de *b*, usando esta ecuación:

La pendiente de una línea recta

$$b = \frac{Y_2 - Y_1}{X_2 - X_1} \qquad [12\text{-}2]$$

$$b = \frac{7 - 5}{2 - 1}$$

$$= \frac{2}{1}$$

$$= 2 \leftarrow \text{Pendiente de la recta}$$

Escritura y uso de la ecuación de una recta

De esta manera, podemos conocer los valores de las constantes numéricas, *a* y *b*, y escribir la ecuación de la recta. La línea de la figura 12-6 puede describirse por la ecuación 12-1, en la que $a = 3$ y $b = 2$. Por tanto,

$$Y = a + bX \qquad [12\text{-}1]$$

y

$$Y = 3 + 2X$$

Usando esta ecuación, podemos determinar el valor correspondiente de la variable dependiente para cualquier valor de *X*. Supongamos que deseamos encontrar el valor de *Y* cuando $X = 7$. La respuesta sería

$$Y = a + bX \qquad [12\text{-}1]$$

$$= 3 + 2(7)$$

$$= 3 + 14$$

$$= 17$$

Relación directa; pendiente positiva

Si sustituye más valores de *X* en la ecuación, observará que *Y* se incrementa al aumentar *X*. Por tanto, la relación entre las variables es *directa* y la pendiente es *positiva*.

Ahora consideremos la línea de la figura 12-7. Vemos que cruza el eje *Y* en 6. Por tanto, sabemos que $a = 6$. Si seleccionamos los dos puntos donde (X_1, Y_1) = (0, 6) y (X_2, Y_2) = (1, 3), encontraremos que la pendiente de la recta es

$$b = \frac{Y_2 - Y_1}{X_2 - X_1} \qquad [12\text{-}2]$$

$$= \frac{3 - 6}{1 - 0}$$
$$= -\frac{3}{1}$$
$$= -3$$

Relación inversa; pendiente negativa

Observe que cuando b es negativa, la recta representa una relación *inversa*, y la pendiente es *negativa* (Y disminuye al aumentar X). Una vez determinados los valores numéricos de a y b, podemos sustituirlos en la ecuación general de la línea recta:

$$Y = a + bX \qquad [12\text{-}1]$$
$$= 6 + (-3)X$$
$$= 6 - 3X$$

Encontrar Y dado X

Suponga que deseamos encontrar el valor de la variable dependiente que corresponde a $X = 2$. Sustituyendo en la ecuación 12-1 obtenemos:

$$Y = 6 - (3)(2)$$
$$= 6 - 6$$
$$= 0$$

Por tanto, cuando $X = 2$, Y debe ser igual a 0. Si consultamos la línea de la figura 12-7, podemos ver que el punto (2, 0) sí está en la recta.

El método de mínimos cuadrados

Ajuste matemático de una recta de regresión

Ahora que hemos visto cómo determinar la ecuación de una línea recta, pensemos cómo calcular una ecuación para una línea dibujada en medio de un conjunto de puntos de un diagrama de dispersión. ¿Cómo podemos "ajustar" una recta matemáticamente si ninguno de los puntos está sobre ella? Para un especialista en estadística, la línea tendrá un "buen ajuste" si *minimiza el error* entre los puntos estimados en la recta y los puntos observados reales que se utilizaron para trazarla.

Introducción de $\hat{Y}$

Antes de proceder, necesitamos introducir un nuevo símbolo. Hasta ahora, hemos utilizado Y para representar los valores individuales de los puntos observados medidos a lo largo del eje Y. Ahora de-

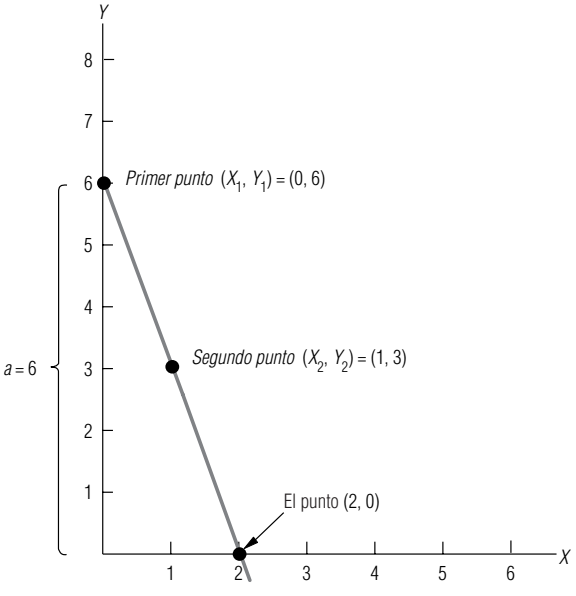


FIGURA 12-7
Línea recta con pendiente negativa

bemos comenzar a usar $\hat{Y}$ *(ye gorro)* para simbolizar los valores individuales de los puntos *estimados*, esto es, aquellos puntos que están en la línea de estimación. En consecuencia, escribiremos la ecuación para la línea de estimación como

La línea de estimación

$$\hat{Y} = a + bX \qquad [12\text{-}3]$$

¿Qué línea se ajusta mejor?

En la figura 12-8, tenemos dos líneas de estimación que se han ajustado al mismo conjunto de tres puntos. Estos tres puntos dados, u observados, se muestran en negro. Se han trazado dos líneas muy diferentes para describir la relación entre las dos variables. Obviamente, necesitamos una forma de decidir cuál de estas líneas nos proporciona un mejor ajuste.

Uso del error total para determinar el mejor ajuste

Una forma en que podemos "medir el error" de nuestra línea de estimación es *sumando* todas las diferencias, o errores, individuales entre los puntos estimados mostrados en círculo y los puntos observados mostrados en negro. En la tabla 12-2, calculamos las diferencias individuales entre las Y correspondientes y $\hat{Y}$, y luego encontramos la suma de estas diferencias.

Tabla 12-2
Suma de errores de las dos líneas de estimación de la figura 12-8

Gráfica (*a*) $Y - \hat{Y}$	Gráfica (*b*) $Y - \hat{Y}$
$8 - 6 = 2$	$8 - 2 = 6$
$1 - 5 = -4$	$1 - 5 = -4$
$6 - 4 = 2$	$6 - 8 = -2$
0 ← **Error total**	**0** ← **Error total**

Tabla 12-3
Suma de los valores absolutos del error de las dos líneas de estimación de la figura 12-8

Gráfica (*a*) $\lvert Y - \hat{Y} \rvert$	Gráfica (*b*) $\lvert Y - \hat{Y} \rvert$
$\lvert 8 - 6 \rvert = 2$	$\lvert 8 - 2 \rvert = 6$
$\lvert 1 - 5 \rvert = 4$	$\lvert 1 - 5 \rvert = 4$
$\lvert 6 - 4 \rvert = 2$	$\lvert 6 - 8 \rvert = 2$
8 ← **Error absoluto total**	**12** ← **Error absoluto total**

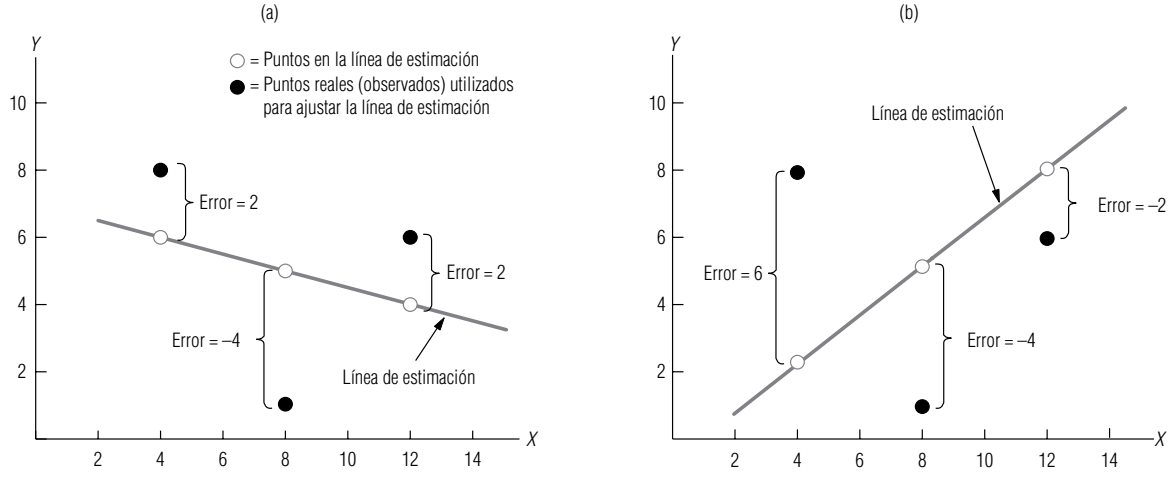


FIGURA 12-8 Dos líneas de estimación diferentes ajustadas a los mismos tres puntos observados; se muestran errores en ambos casos

Un rápido examen visual de las dos líneas de estimación de la figura 12-8 revela que la línea de la gráfica (a) se ajusta a los tres puntos de datos mejor que la línea de la gráfica (b).* Sin embargo, nuestro proceso de suma de las diferencias individuales de la tabla 12-2 indica que ambas líneas describen los datos igualmente bien (el error total en ambos casos es cero). Por tanto, debemos concluir que el proceso de suma de las diferencias individuales para calcular el error no es una forma confiable de juzgar la bondad de ajuste de una línea de estimación.

Uso del valor absoluto del error para medir el mejor ajuste

El problema al sumar los errores individuales es el efecto de cancelación de los valores positivos y negativos. De esto, podríamos deducir que el criterio adecuado para juzgar la bondad del ajuste sería sumar los valores absolutos (los valores sin los signos algebraicos) de cada error. Hemos hecho esto en la tabla 12-3. (El símbolo del valor absoluto son dos líneas verticales paralelas, ||.) Como el error absoluto en la gráfica (a) es menor que el error absoluto en la gráfica (b), dado que buscamos el "mínimo error absoluto", confirmamos nuestra impresión intuitiva de que la línea de estimación de la gráfica (a) es el mejor ajuste.

Con base en este éxito, podríamos concluir que la minimización de la suma de los valores absolutos de los errores es el mejor criterio para encontrar un buen ajuste. Pero antes de sentirnos demasiado cómodos con él, debemos examinar una situación distinta.

La figura 12-9 nuevamente presenta dos diagramas de dispersión idénticos con dos líneas de estimación diferentes ajustadas a los tres datos puntuales. En la tabla 12-4, sumamos los valores absolutos de los errores y encontramos que la línea de estimación de la gráfica (a) es un mejor ajuste que la de la gráfica (b). Intuitivamente, sin embargo, pareciera que la línea de la gráfica (b) es la mejor línea de ajuste, porque se ha movido verticalmente para tomar el punto medio en consideración. La gráfica (a), por otra parte, parece ignorar completamente el punto medio. Así que tal vez descartaríamos este segundo criterio para encontrar el mejor ajuste. ¿Por qué? **La suma de los valores absolutos no hace hincapié en la *magnitud* del error.**

Dar más peso a los puntos más lejanos; elevar el error al cuadrado

Parece razonable que mientras más lejos esté un punto de la línea de estimación, más serio es el error. Preferiríamos tener varios errores absolutos pequeños que uno grande, como vimos en el ejemplo anterior. **En efecto, deseamos encontrar una forma de "penalizar" errores absolutos grandes, para poder evitarlos. Podemos lograr esto si elevamos *al cuadrado* los errores individuales antes de sumarlos.** Los cuadrados de cada término logran dos objetivos:

1. Magnifica, o penaliza, los errores más grandes.
2. Cancela el efecto de los valores positivos y negativos (un error negativo al cuadrado sigue siendo positivo).

Uso de mínimos cuadrados como una medida del mejor ajuste

Como estamos buscando la línea de estimación que minimiza la suma de los cuadrados de los errores, a esto le llamamos *método de mínimos cuadrados*.

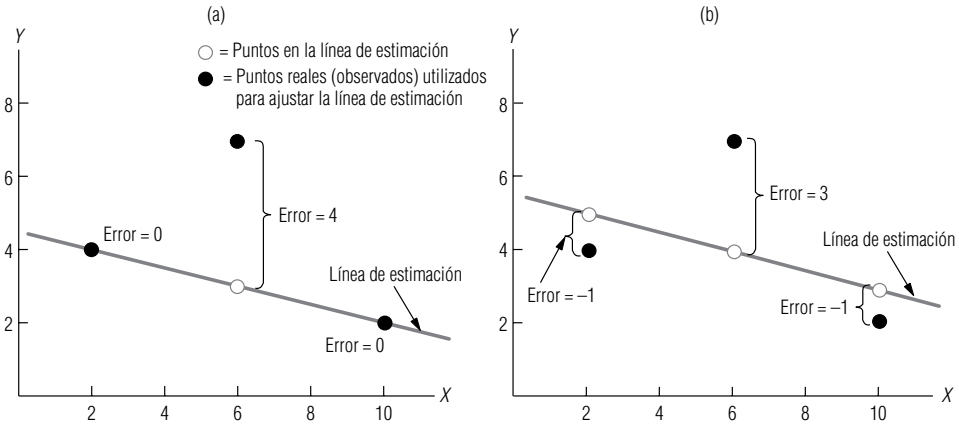


FIGURA 12-9

Dos líneas de estimación diferentes ajustadas a los mismos puntos observados; se muestran errores en ambos casos

* Podemos razonar que esto es así al observar que mientras ambas líneas de estimación se separan del segundo y tercer punto (de izquierda a derecha) una distancia igual, la línea de la gráfica (a) se separa del primer punto una distancia mucho menor que la línea de la gráfica (b).

Tabla 12-4

Suma de los valores absolutos de los errores de las dos líneas de estimación de la figura 12-9

Gráfica (*a*) $\lvert Y - \hat{Y} \rvert$	Gráfica (*b*) $\lvert Y - \hat{Y} \rvert$
$\lvert 4 - 4 \rvert = 0$	$\lvert 4 - 5 \rvert = 1$
$\lvert 7 - 3 \rvert = 4$	$\lvert 7 - 4 \rvert = 3$
$\lvert 2 - 2 \rvert = 0$	$\lvert 2 - 3 \rvert = 1$
4 ← Error absoluto total	**5** ← Error absoluto total

Tabla 12-5

Aplicación del criterio de mínimos cuadrados a las líneas de estimación

Gráfica (*a*) $(Y - \hat{Y})^2$	Gráfica (*b*) $(Y - \hat{Y})^2$
$(4 - 4)^2 = (0)^2 = 0$	$(4 - 5)^2 = (-1)^2 = 1$
$(7 - 3)^2 = (4)^2 = 16$	$(7 - 4)^2 = (3)^2 = 9$
$(2 - 2)^2 = (0)^2 = 0$	$(2 - 3)^2 = (-1)^2 = 1$
16 ← Suma de cuadrados	**11** ← Suma de cuadrados

Apliquemos el criterio de mínimos cuadrados al problema de la figura 12-9. Una vez que organizamos los datos y sumamos los cuadrados de la tabla 12-5, vemos que, tal como lo pensamos, la línea de estimación en la gráfica (b) es el mejor ajuste.

Cómo encontrar matemáticamente la recta de mínimos cuadrados que mejor se ajusta

Usando el criterio de los mínimos cuadrados, podemos determinar si una línea de estimación es mejor ajuste que otro. Pero para un conjunto de puntos a través de los cuales podríamos trazar un número infinito de líneas de estimación, ¿cómo podemos saber cuándo hemos encontrado *la recta del mejor ajuste*?

Los estadísticos han desarrollado dos ecuaciones que podemos utilizar para encontrar la pendiente y la ordenada Y de la recta de regresión de mejor ajuste. La primera fórmula calcula la pendiente:

Pendiente de la recta de regresión de mínimos cuadrados

Pendiente de la recta de regresión de mejor ajuste

$$b = \frac{\Sigma XY - n\overline{X}\,\overline{Y}}{\Sigma X^2 - n\overline{X}^2} \qquad [12\text{-}4]$$

donde,

- b = pendiente de la línea de estimación de mejor ajuste
- X = valores de la variable independiente
- Y = valores de la variable dependiente
- $\overline{X}$ = media de los valores de la variable independiente
- $\overline{Y}$ = media de los valores de la variable dependiente
- n = número de puntos (es decir, el número de pares de valores de las variables independiente y dependiente)

La segunda fórmula calcula la ordenada Y de la recta cuya pendiente calculamos usando la ecuación 12-4:

Ordenada de la recta de regresión de mínimos cuadrados

Ordenada *Y* de la recta de regresión de mejor ajuste

$$a = \overline{Y} - b\overline{X} \qquad [12\text{-}5]$$

donde,

- a = ordenada Y
- b = pendiente de la ecuación 12-4

- $\overline{Y}$ = media de los valores de la variable dependiente
- $\overline{X}$ = media de los valores de la variable independiente

Con estas dos ecuaciones, podemos encontrar la recta de regresión de mejor ajuste para cualquier conjunto de puntos para dos variables.

Uso del método de mínimos cuadrados en dos problemas

Suponga que la directora del Departamento de Salubridad de Chapel Hill está interesada en la relación que existe entre la antigüedad de un camión de basura y los gastos anuales de reparación que debe esperar. Con el fin de determinar esta relación, la directora ha reunido información de cuatro de los camiones de la ciudad (tabla 12-6).

Ejemplo del método de mínimos cuadrados

El primer paso para calcular la recta de regresión de este problema es organizar los datos como se resumen en la tabla 12-7. Esto nos permite sustituirlos directamente en las ecuaciones 12-4 y 12-5 para encontrar la pendiente y la ordenada Y de la recta de regresión de mejor ajuste.

Con la información de la tabla 12-7, podemos usar las ecuaciones para la pendiente (ecuación 12-4) y para la ordenada Y (ecuación 12-5) con el fin de encontrar las constantes numéricas para la recta de regresión. La pendiente es:

Búsqueda del valor de *b*

$$\begin{aligned} b &= \frac{\Sigma XY - n\overline{X}\,\overline{Y}}{\Sigma X^2 - n\overline{X}^2} \\ &= \frac{78 - (4)(3)(6)}{44 - (4)(3)^2} \\ &= \frac{78 - 72}{44 - 36} \\ &= \frac{6}{8} \\ &= 0.75 \leftarrow \text{Pendiente de la línea} \end{aligned}$$

Y la ordenada Y es:

Búsqueda del valor de *a*

$$\begin{aligned} a &= \overline{Y} - b\overline{X} \qquad [12\text{-}5] \\ &= 6 - (0.75)(3) \\ &= 6 - 2.25 \\ &= 3.75 \leftarrow \text{Ordenada } Y \end{aligned}$$

Determinación de la ecuación de estimación

Ahora, para obtener la ecuación de estimación que describe la relación entre la antigüedad de un camión y sus gastos anuales de reparación, podemos sustituir los valores de a y b en la ecuación general para una línea recta:

$$\begin{aligned} \hat{Y} &= a + bX \qquad [12\text{-}3] \\ &= 3.75 + 0.75X \end{aligned}$$

Tabla 12-6
Gastos anuales de reparación de camiones

Número del camión	Antigüedad del camión en años (X)	Gastos de rep. durante el último año en cientos de dólares (Y)
101	5	7
102	3	7
103	3	6
104	1	4

Tabla 12-7

Cálculo de los datos para las ecuaciones 12-4 y 12-5

Camiones (n = 4) (1)	Antigüedad (X) (2)	Gastos de reparación (Y) (3)	XY (2) × (3)	X^2 $(2)^2$
101	5	7	35	25
102	3	7	21	9
103	3	6	18	9
104	1	4	4	1
	$\Sigma X = 12$	$\Sigma Y = 24$	$\Sigma XY = 78$	$\Sigma X^2 = 44$

$$\bar{X} = \frac{\Sigma X}{n}$$

$$= \frac{12}{4}$$

$= 3 \leftarrow$ Media de los valores de la variable independiente

$$\bar{Y} = \frac{\Sigma X}{n}$$

$$= \frac{24}{6}$$

$= 6 \leftarrow$ Media de los valores de la variable dependiente

Tabla 12-8

Relación anual entre investigación, desarrollo y ganancias

Año	Millones de dólares gastados en investigación y desarrollo (*X*)	Ganancia anual (millones de dólares) (*Y*)
1995	5	31
1994	11	40
1993	4	30
1992	5	34
1991	3	25
1990	2	20

Uso de la ecuación de estimación

Utilizando esta ecuación de estimación (que podríamos graficar como una recta de regresión si así lo deseáramos), la directora del Departamento de Salubridad puede estimar los gastos anuales de reparación, dada la antigüedad de su equipo. Si, por ejemplo, la ciudad tiene un camión de 4 años de antigüedad, la directora podría usar la ecuación para predecir los gastos anuales de reparación para este camión de la siguiente manera:

$$\hat{Y} = 3.75 + 0.75(4)$$

$$= 3.75 + 3$$

$= 6.75 \leftarrow$ Gastos anuales de reparación esperados de \$675.00

Así, se calcularía que la ciudad gasta aproximadamente \$675 al año en reparaciones de un camión de 4 años de antigüedad.

Otro ejemplo

Ahora podemos resolver el problema del inicio del capítulo, referente a la relación entre el dinero gastado en investigación y desarrollo y las ganancias anuales de la compañía química. La tabla 12-8 presenta la información de los 6 años anteriores. Con esto, podemos determinar la ecuación de regresión que describe la relación.

Nuevamente, podemos facilitar la recolección de la información necesaria si realizamos los cálculos de la tabla 12-9.

Con esta información, estamos listos para encontrar las constantes numéricas *a* y *b* para la ecuación de estimación. El valor de *b* es:

Cálculo de *b*

$$b = \frac{\Sigma XY - n\bar{X}\,\bar{Y}}{\Sigma X^2 - n\bar{X}^2} \qquad [12\text{-}4]$$

$$= \frac{1{,}000 - (6)(5)(30)}{200 - (6)(5)^2}$$

$$= \frac{1{,}000 - 900}{200 - 150}$$

$$= \frac{100}{50}$$

$$= 2 \leftarrow \text{Pendiente de la recta}$$

Y el valor de *a* es:

Cálculo de *a*

$$a = \bar{Y} - b\bar{X} \qquad [12\text{-}5]$$

$$= 30 - (2)(5)$$

$$= 30 - 10$$

$$= 20 \leftarrow \text{Ordenada } Y$$

Determinación de la ecuación de estimación

Entonces podemos sustituir estos valores de *a* y *b* en la ecuación 12-3 y obtener:

$$\hat{Y} = a + bX \qquad [12\text{-}3]$$

$$= 20 + 2X$$

Uso de la ecuación de estimación para pronosticar

Al utilizar esta ecuación de estimación, el vicepresidente de investigación y desarrollo puede predecir las ganancias futuras anuales a partir de la cantidad presupuestada para ID. Si la compañía gastó 8 millones de dólares para ID en 1996, entonces debió ganar aproximadamente 36 millones de dólares durante ese año:

Tabla 12-9

Cálculo de los datos para las ecuaciones 12-4 y 12-5

Año ($n = 6$)	Gastos de ID (X)	Ganancias anuales (Y)	XY	X^2
1995	5	31	155	25
1994	11	40	440	121
1993	4	30	120	16
1992	5	34	170	25
1991	3	25	75	9
1990	2	20	40	4
	$\Sigma X = 30$	$\Sigma Y = 180$	$\Sigma XY = 1{,}000$	$\Sigma X^2 = 200$

$$\bar{X} = \frac{\Sigma X}{n} \qquad [3\text{-}2]$$

$$= \frac{30}{6}$$

$$= 5 \leftarrow \text{Media de los valores de la variable independiente}$$

$$\bar{Y} = \frac{\Sigma Y}{n} \qquad [3\text{-}2]$$

$$= \frac{180}{6}$$

$$= 30 \leftarrow \text{Media de los valores de la variable dependiente}$$

$$\hat{Y} = 20 + 2(8)$$
$$= 20 + 16$$
$$= 36 \leftarrow \text{Ganancia anual esperada (millones de dólares)}$$

Deficiencia de la ecuación de estimación para predecir

Las ecuaciones de estimación no son pronosticadores perfectos. En la figura 12-10, que grafica los puntos encontrados en la tabla 12-8, la estimación de 36 millones de ganancia para 1996 es sólo eso, una estimación. Aun así, la regresión sí nos da una idea de qué esperar para el siguiente año.

Verificación de la ecuación de estimación

Una forma de verificar la ecuación de estimación

Ahora que sabemos cómo calcular la línea de regresión, podemos aprender cómo verificar nuestro trabajo. Una forma burda de verificar la exactitud de la ecuación de estimación es examinar la gráfica de los puntos de la muestra. Como podemos ver del problema anterior, la línea de regresión de la figura 12-10 parece seguir la trayectoria descrita por los puntos de la muestra.

Otra forma de verificar la ecuación de estimación

Un método más sofisticado surge de una de las propiedades matemáticas de una recta ajustada por el método de mínimos cuadrados, es decir, los errores individuales positivos y negativos deben sumar cero. Usando la información de la tabla 12-9, verifique que la suma de los errores en el último problema sea igual a cero. Esto se hace en la tabla 12-10.

Como la suma de los errores de la tabla 12-10 sí es igual a cero, y puesto que la línea de regresión parece "ajustarse" a los puntos de la figura 12-10, podemos estar razonablemente seguros de que no hemos cometido errores matemáticos serios al determinar la ecuación de estimación para este problema.

Tabla 12-10

Cálculo de la suma de los errores individuales de la tabla 12-9

Y		$\hat{Y}$ (es decir, $20 + 2X$)		Error individual
31	−	[20 + (2)(5)]	=	1
40	−	[20 + (2)(11)]	=	−2
30	−	[20 + (2)(4)]	=	2
34	−	[20 + (2)(5)]	=	4
25	−	[20 + (2)(3)]	=	−1
20	−	[20 + (2)(2)]	=	−4
				0 ← Error total

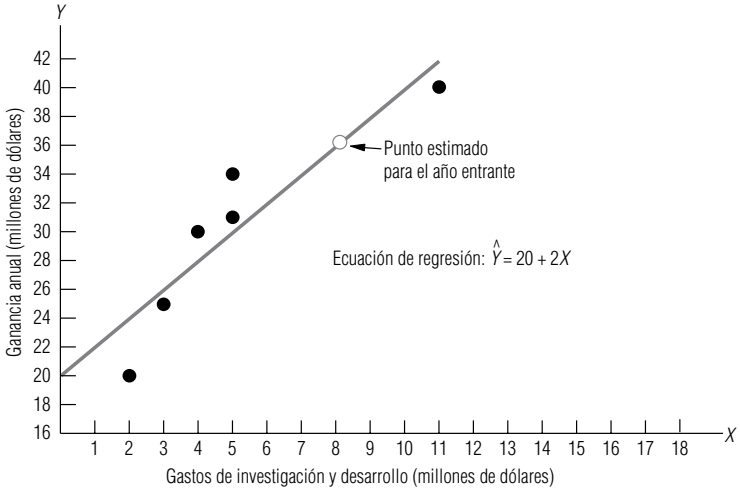


FIGURA 12-10

Dispersión de puntos alrededor de la línea de regresión

El error estándar de la estimación

Medición de la confiabilidad de la ecuación de estimación

El siguiente proceso que debemos aprender en nuestro estudio del análisis de regresión es cómo medir la confiabilidad de la ecuación de estimación desarrollada. Aludimos a este tema cuando introdujimos los diagramas de dispersión; en ese punto, nos dimos cuenta intuitivamente de que una línea será más exacta como estimador cuando los datos puntuales caen cerca de la línea [como en la gráfica (a) de la figura 12-11] que cuando los puntos están alejados de la línea [como en la gráfica (b) de la figura 12-1l].

Definición y uso del error estándar de la estimación

Para medir la confiabilidad de la ecuación de estimación, los especialistas en estadística han desarrollado el *error estándar de la estimación.* Este error estándar se simboliza por s_e y es similar a la desviación estándar (que examinamos por primera vez en el capítulo 3), en cuanto a que ambas son medidas de dispersión. Recordará que la desviación estándar se utiliza para medir la dispersión de un conjunto de observaciones respecto a la media. **El error estándar de la estimación, por otra parte, mide la variabilidad, o dispersión, de los valores observados alrededor de la recta de regresión.** Aun así, verá la similitud entre el error estándar de la estimación y la desviación estándar si compara la ecuación 12-6, que define el error estándar de la estimación, con la ecuación 3-18, que define la desviación estándar:

Ecuación para calcular el error estándar de la estimación

Error estándar de la estimación

$$s_e = \sqrt{\frac{\Sigma(Y - \hat{Y})^2}{n - 2}} \qquad [12\text{-}6]$$

donde,

- Y = valores de la variable dependiente
- $\hat{Y}$ = valores estimados con la ecuación de estimación que corresponden a cada valor de Y
- n = número de puntos utilizados para ajustar la línea de regresión

$n - 2$ es el divisor en la ecuación 12-6

Observe que, en la ecuación 12-6, la suma de las desviaciones al cuadrado se divide entre $n - 2$ y no entre n. Esto sucede porque perdimos dos grados de libertad al estimar la recta de regresión. Podemos razonar que, dado que los valores de a y b se obtuvieron de una muestra de datos puntuales, perdemos dos grados de libertad cuando usamos estos puntos para estimar la recta de regresión.

Ahora, no referiremos de nuevo al ejemplo anterior de la directora del Departamento de Salubridad que relacionaba la antigüedad de sus camiones con la cantidad de reparaciones anuales. Encontramos que la ecuación de estimación en esa situación era:

$$\hat{Y} = 3.75 + 0.75X$$

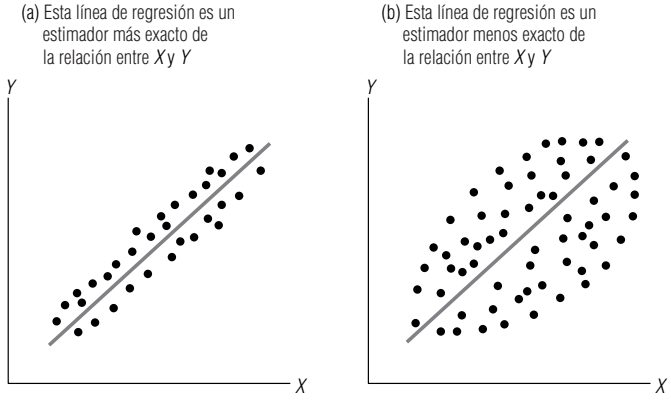


FIGURA 12-11

Grados contrastantes de dispersión de datos puntuales y el efecto resultante en la precisión de la recta de regresión

donde X es la antigüedad del camión y $\hat{Y}$ la cantidad estimada de reparaciones anuales (en cientos de dólares).

Cálculo del error estándar de la estimación

Para calcular s_e para este problema, primero debemos determinar el valor de $\sum(Y - \hat{Y})^2$, esto es, el numerador de la ecuación 12-6. Hicimos esto en la tabla 12-11, usando $(3.75 + 0.75X)$ para $\hat{Y}$, siempre que fue necesario. Como $\sum(Y - \hat{Y})^2$ es igual a 1.50, podemos usar la ecuación 12-6 para encontrar el error estándar de la estimación:

$$s_e = \sqrt{\frac{\Sigma(Y - \hat{Y}^2)}{n - 2}} \qquad [12\text{-}6]$$

$$= \sqrt{\frac{1.50}{4 - 2}}$$

$$= \sqrt{0.75}$$

$$= 0.866 \leftarrow \text{Error estándar de la estimación de \$86.60}$$

Uso de un método abreviado para calcular el error estándar de la estimación

Para usar la ecuación 12-6, debemos hacer la tediosa serie de cálculos descritos en la tabla 12-11. Para cada valor de Y, debemos calcular el valor correspondiente de $\hat{Y}$. Entonces debemos sustituir estos valores en la expresión $\sum(Y - \hat{Y})^2$.

Afortunadamente, podemos eliminar algunos pasos de esta tarea al usar el camino corto proporcionado por la ecuación 12-7, esto es:

Una forma más rápida de calcular s_e

Método abreviado para encontrar el error estándar de la estimación

$$s_e = \sqrt{\frac{\Sigma Y^2 - a\Sigma Y - b\Sigma XY}{n - 2}} \qquad [12\text{-}7]$$

donde,

- X = valores de la variable independiente
- Y = valores de la variable dependiente
- a = ordenada Y de la ecuación 12-5
- b = pendiente de la ecuación de estimación de la ecuación 12-4
- n = número de puntos

Esta ecuación es un atajo, porque al organizar primero los datos de este problema para calcular la pendiente y la ordenada Y (tabla 12-7), determinamos cada valor que necesitamos para la ecuación

Tabla 12-11

Cálculo del numerador de la fracción en la ecuación 12-6

X (1)	Y (2)	$\hat{Y}$ (es decir, $3.75 + 0.75X$) (3)	Error individual $(Y - \hat{Y})$ (2) − (3)	$(Y - \hat{Y})^2$ [(2) − (3)]^2
5	7	$3.75 + (0.75)(5)$	$7 - 7.5 = -0.5$	0.25
3	7	$3.75 + (0.75)(3)$	$7 - 6.0 = 1.0$	1.00
3	6	$3.75 + (0.75)(3)$	$6 - 6.0 = 0.0$	0.00
1	4	$3.75 + (0.75)(1)$	$4 - 4.5 = -0.5$	0.25
				$\Sigma(Y - \hat{Y})^2 = 1.50$ ← Suma de los cuadrados de los errores

Tabla 12-12
Calculo de los datos para la ecuación 12-7

Camiones $n = 4$ (1)	Antigüedad (X) (2)	Gastos de reparación (Y) (3)	XY (2) × (3)	X^2 $(2)^2$	Y^2 $(3)^2$
101	5	7	35	25	49
102	3	7	21	9	49
103	3	6	18	9	36
104	1	4	4	1	16
	$\Sigma X = 12$	$\Sigma Y = 24$	$\Sigma XY = 78$	$\Sigma X2 = 44$	$\Sigma Y2 = 150$

12-7, excepto uno: el valor de $\sum Y^2$. La tabla 12-12 es una repetición de la tabla 12-7, añadiendo la columna Y^2.

Ahora podemos consultar la tabla 12-12 y nuestros cálculos anteriores de a y b, con el fin de calcular s_e usando el método abreviado:

$$s_e = \sqrt{\frac{\sum Y^2 - a\sum Y - b\sum XY}{n-2}} \qquad [12\text{-}7]$$

$$= \sqrt{\frac{150 - (3.75)(24) - (0.75)(78)}{4-2}}$$

$$= \sqrt{\frac{150 - 90 - 58.5}{2}}$$

$$= \sqrt{0.75}$$

$$= 0.866 \leftarrow \text{Error estándar de \$86.60}$$

Éste resultado es igual al obtenido usando la ecuación 12-6, ¡pero piense en cuántos pasos nos ahorramos!

Interpretación del error estándar de la estimación

Interpretación y uso del error estándar de la estimación

Como ocurría en el caso de la desviación estándar, mientras más grande sea el error estándar de la estimación, mayor será la dispersión de los puntos alrededor de la línea de regresión. De manera inversa, si $s_e = 0$, esperamos que la ecuación de estimación sea un estimador "perfecto" de la variable dependiente. En ese caso, todos los puntos caerían directamente sobre la línea de regresión y no habría puntos dispersos alrededor.

Uso de s_e para formar límites alrededor de la línea de regresión

Usaremos el error estándar de la estimación como una herramienta, de la misma forma que podemos usar la desviación estándar. Esto es, suponiendo que los puntos observados siguen una distribución normal alrededor de la recta de regresión, podemos esperar encontrar el 68% de los puntos dentro de $\pm 1 s_e$ (o más menos 1 error estándar de la estimación), el 95.5% de los puntos dentro de $\pm 2 s_e$ y el 99.7% de los puntos dentro de $\pm 3 s_e$. La figura 12-12 ilustra estos "límites" alrededor de la línea de regresión. **Otra cosa que debemos observar en la figura 12-12 es que el error estándar de la estimación se mide a lo largo del eje *Y*, y no perpendicularmente desde la recta de regresión.**

Suposiciones para usar s_e

En este punto, debemos establecer las suposiciones necesarias, ya que pronto haremos algunas afirmaciones probabilísticas. Específicamente:

1. Los valores observados para Y tienen distribución normal alrededor de cada valor estimado de $\hat{Y}$.
2. La varianza de las distribuciones alrededor de cada valor posible de $\hat{Y}$ es la misma.

Si esta segunda suposición no fuera cierta, entonces el error estándar en un punto de la recta de regresión podría diferir del error estándar en otro punto.

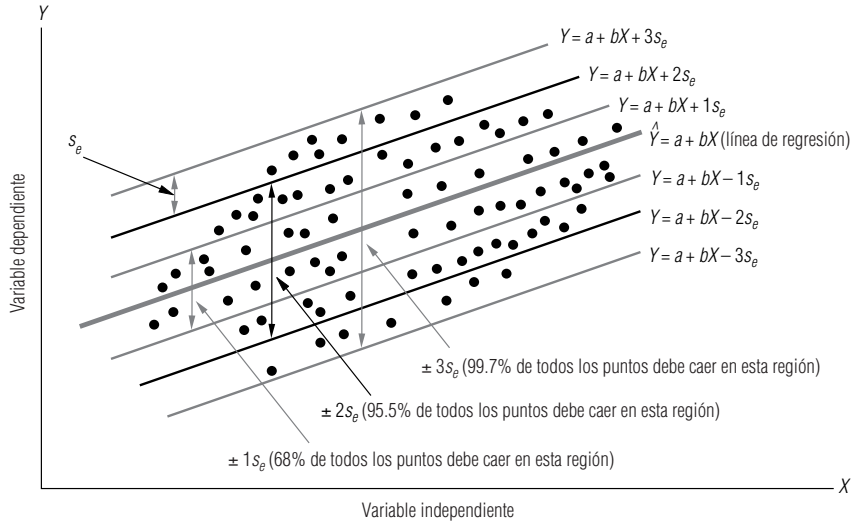


FIGURA 12-12

límites alrededor de la línea de regresión de $\pm 1s_e$ $\pm 2s_e$ y $\pm 3s_e$

Intervalos de confianza para la estimación (o el valor esperado)

Utilización de s_e para generar intervalos de confianza

Podemos concebir al error estándar de la estimación como la herramienta estadística que podemos usar para hacer afirmaciones de probabilidad acerca del intervalo alrededor del valor estimado de $\hat{Y}$, dentro del cual cae el valor real de Y. En la figura 12-12 podemos ver, por ejemplo, que hay una seguridad del 95.5% de que el valor real de Y caerá dentro de dos errores estándar del valor estimado de $\hat{Y}$. Llamamos a estos intervalos alrededor de la $\hat{Y}$ estimada, *intervalos de confianza para la estimación*. Tienen la misma función que los intervalos de confianza en el capítulo 7.

Ahora, aplicando el concepto de intervalos de confianza para la estimación al problema de la directora del Departamento de Salubridad, sabemos que la ecuación de estimación usada para predecir el gasto anual de reparación es:

$$\hat{Y} = 3.75 + 0.75X$$

Aplicación de los intervalos de confianza para la estimación (o valor esperado)

Y sabemos que si el departamento tiene un camión de cuatro años de antigüedad, predecimos que tendrá un gasto de reparaciones anuales de $675:

$$\begin{aligned}\hat{Y} &= 3.75 + 0.75(4)\\ &= 3.75 + 3.00\\ &= 6.75 \leftarrow \text{Gasto anual de reparaciones esperado de \$675}\end{aligned}$$

Intervalo de confianza para la estimación de un error estándar

Por último, recordará que calculamos el error estándar de la estimación como $s_e = 0.866$ ($86.60). Ahora podemos combinar estas dos piezas de información y decir que estamos seguros aproximadamente el 68% del tiempo, de que el gasto real de reparaciones estará dentro de ± 1 error estándar de la estimación de $\hat{Y}$. Podemos calcular los límites superior e inferior de este intervalo de confianza para el gasto de reparación de la siguiente manera:

$$\begin{aligned}\hat{Y} + 1s_e &= \$675 + (1)(\$86.60)\\ &= \$761.40 \leftarrow \text{Límite superior del intervalo de predicción}\end{aligned}$$

y

$$\begin{aligned}\hat{Y} - 1s_e &= \$675 - (1)(\$86.60)\\ &= \$588.40 \leftarrow \text{Límite inferior del intervalo de predicción}\end{aligned}$$

Intervalo de confianza para la estimación de dos errores estándar

Si, en lugar de esto, decimos que estamos seguros aproximadamente el 95.5% del tiempo de que el gasto real de reparaciones estará dentro de ± 2 errores estándar de la estimación de $\hat{Y}$, podríamos calcular los límites de este nuevo intervalo de confianza de la siguiente manera:

$$\hat{Y} + 2s_e = \$675 + (2)(\$86.60)$$
$$= \$848.20 \leftarrow \text{Límite superior}$$

y

$$\hat{Y} - 2s_e = \$675 - (2)(\$86.60)$$
$$= \$501.80 \leftarrow \text{Límite inferior}$$

n es demasiado pequeña para usar la distribución normal

Recuerde que los estadísticos aplican los intervalos de confianza para la estimación basados en la distribución normal (el 68% para $1s_e$, el 95.5% para $2s_e$ y el 99.7% para $3s_e$) *sólo* para muestras grandes, esto es, cuando $n > 30$. En este problema, nuestro tamaño de muestra es demasiado pequeño ($n = 4$). Por tanto, *nuestras conclusiones son inexactas*. Pero de todos modos el método que hemos utilizado demuestra el principio involucrado en los intervalos de confianza para la estimación.

Utilización de la distribución *t* para intervalos de confianza para la estimación

Si deseamos evitar inexactitudes ocasionadas por el tamaño de la muestra, necesitamos usar la distribución *t*. Recuerde que esta distribución *t* es apropiada cuando *n* es menor que 30 y la desviación estándar de la población no se conoce. Estas dos condiciones, se cumplen puesto que $n = 4$, y s_e es una estimación y no la desviación estándar conocida de la población.

Un ejemplo del uso de la distribución *t* para calcular intervalos de confianza para la estimación

Ahora suponga que la directora del Departamento de Salubridad desea tener una seguridad aproximada del 90% de que los gastos anuales de reparación caerán en el intervalo de la estimación. ¿Cómo calculamos este intervalo? Como la tabla de distribución *t* se concentra en la probabilidad de que el parámetro que estamos estimando caerá fuera del intervalo de predicción, necesitamos consultar la tabla 2 del apéndice en la columna de 100% − 90% = 10%. Una vez localizada la columna, buscamos el renglón para 2 grados de libertad; porque $n = 4$ y sabemos que perdemos 2 grados de libertad (al estimar los valores de *a* y *b*), entonces $n - 2 = 2$. Encontraremos que el valor apropiado *t* es 2.920.

Ahora, usando este valor de *t*, podemos hacer un cálculo más exacto de los límites del intervalo de la estimación, de la siguiente manera:

$$\hat{Y} + t(s_e) = \$675 + (2.920)(\$86.60)$$
$$= \$675 + \$252.87$$
$$= \$927.87 \leftarrow \text{Límite superior}$$

y

$$\hat{Y} - t(s_e) = \$675 - (2.920)(\$86.60)$$
$$= \$675 - \$252.87$$
$$= \$422.13 \leftarrow \text{Límite inferior}$$

Así, la directora puede estar 90% segura de que los gastos anuales de reparación de un camión de cuatro años de antigüedad estarán entre $422.13 y $927.87.

Debemos resaltar que estos intervalos de la estimación es lo que *se espera que ocurra*. De hecho, los especialistas en estadística pueden calcular el error estándar exacto para calcular intervalos de estimación s_p, usando la fórmula:

$$s_p = s_e \sqrt{1 + \frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\Sigma X^2 - n\bar{X}^2}}$$

donde X_0 es el valor específico de *X* para el que deseamos predecir el valor de *Y*.

Observe que si usamos esta fórmula, s_p será diferente para cada valor de X_0. En particular, si X_0 está *lejos* de $\bar{X}$, entonces s_p será grande, porque $(X - \bar{X})^2$ será grande. Si, por otra parte, X_0 está cerca de *X*, y *n* es moderadamente grande (mayor que 10), entonces s_p estará cerca de s_e. Esto sucede porque $1/n$ es pequeño y $(X_0 - \bar{X})^2$ también lo es. Por tanto, el valor dentro de la raíz cuadrada es cercano a 1, la raíz cuadrada es aún más cercana a 1 y s_p, estará muy cerca de s_e. Esto justifica nuestra utilización de s_e para calcular intervalos de estimación aproximados.

SUGERENCIAS Y SUPOSICIONES

Sugerencia: antes de dedicar tiempo al cálculo de una recta de regresión para un conjunto de datos, tiene sentido dibujar un diagrama de dispersión para esos puntos. Esto permite investigar los puntos distantes porque quizá algunos datos no representen el problema que se desea resolver. Por ejemplo, el gerente de una cadena de restaurantes cerca de la universidad, quien quiere examinar la hipótesis de que las ventas a la hora del almuerzo bajan en días calurosos, puede encontrar que los datos reunidos durante vacaciones y días festivos distorsionan una regresión que de otra manera sería útil. No pierda de vista que es peligroso escoger entre los datos sólo porque se "ajusten" o no a una idea preconcebida de cuál debe ser la conclusión. En el análisis de regresión, la selección cuidadosa y el uso consistente de la mejor base de datos lleva a la ecuación de estimación más valiosa.

Ejercicios 12.2

Ejercicios de autoevaluación

EA 12-2 Para el siguiente conjunto de datos:

a) dibuje un diagrama de dispersión,
b) desarrolle la ecuación de estimación que mejor describa los datos,
c) pronostique Y para $X = 10, 15, 20$.

X	13	16	14	11	17	9	13	17	18	12
Y	6.2	8.6	7.2	4.5	9.0	3.5	6.5	9.3	9.5	5.7

EA 12-3 A menudo, quienes hacen la contabilidad de costos estiman los gastos generales con base en el nivel de producción. En Standard Knitting Co. han reunido información acerca de los gastos generales y las unidades producidas en diferentes plantas, y ahora desean estimar una ecuación de regresión para predecir los gastos generales futuros.

Gastos generales	191	170	272	155	280	173	234	116	153	178
Unidades	40	42	53	35	56	39	48	30	37	40

a) Desarrolle una ecuación de regresión para contabilidad de costos.
b) Pronostique los gastos generales cuando se producen 50 unidades.
c) Calcule el error estándar de la estimación.

Conceptos básicos

■ **12-13** Para los siguientes datos:

a) trace un diagrama de dispersión,
b) desarrolle la ecuación de estimación que mejor describa los datos,
c) pronostique Y para $X = 6, 13.4, 20.5$.

X	2.7	4.8	5.6	18.4	19.6	21.5	18.7	14.3
Y	16.66	16.92	22.3	71.8	80.88	81.4	77.46	48.7

X	11.6	10.9	18.4	19.7	12.3	6.8	13.8
Y	50.48	47.82	71.5	81.26	50.1	39.4	52.8

■ **12-14** Usando los datos dados a continuación,

a) trace el diagrama de dispersión,
b) desarrolle la ecuación de estimación que mejor describa los datos,
c) pronostique Y para $X = 5, 6, 7$.

X	16	6	10	5	12	14
Y	−4.4	8.0	2.1	8.7	0.1	−2.9

■ **12-15** Dado el siguiente conjunto de datos:

a) encuentre la línea de mejor ajuste,

b) calcule el error estándar de la estimación,
c) encuentre un intervalo de la estimación aproximada (con el 95% de nivel de confianza) para la variable dependiente dado que X es 44.

X	56	48	42	58	40	39	50
Y	45	38.5	34.5	46.1	33.3	32.1	40.4

Aplicaciones

■ **12-16** Las ventas de línea blanca varían según el estado del mercado de casas nuevas: cuando las ventas de casas nuevas son buenas, también lo son las de lavaplatos, lavadoras de ropa, secadoras y refrigeradores. Una asociación de comercio compiló los siguientes datos históricos (en miles de unidades) de las ventas de línea blanca y la construcción de casas.

Construcción de casas (miles)	Ventas de línea blanca (miles)
2.0	5.0
2.5	5.5
3.2	6.0
3.6	7.0
3.3	7.2
4.0	7.7
4.2	8.4
4.6	9.0
4.8	9.7
5.0	10.0

a) Desarrolle una ecuación para la relación entre las ventas de línea blanca (en miles) y la construcción de casas (en miles).
b) Interprete la pendiente de la recta de regresión.
c) Calcule e interprete el error estándar de la estimación.
d) La construcción de casas durante el año próximo puede ser mayor que el intervalo registrado; se han pronosticado estimaciones hasta de 8.0 millones de unidades. Calcule un intervalo de predicción de 90% de confianza para las ventas de línea blanca, con base en los datos anteriores y el nuevo pronóstico de construcción de casas.

■ **12-17** Durante partidos recientes de tenis, Diane ha observado que sus lanzamientos no han sido eficaces, pues sus oponentes le han regresado algunos de ellos. Algunas de las personas con las que juega son bastante altas, así que se pregunta si la estatura de su contrincante podría explicar el número de lanzamientos no regresados durante un partido. Los siguientes datos se sacaron de cinco partidos recientes.

Estatura del oponente (*H*)	Lanzamientos no regresados (*L*)
5.0	9
5.5	6
6.0	3
6.5	0
5.0	7

a) ¿Cuál es la variable dependiente?
b) ¿Cuál es la ecuación de estimación de mínimos cuadrados para estos datos?
c) ¿Cuál es su mejor estimación del número de lanzamientos no regresados en su partido de mañana con un oponente de 5.9 pies de estatura?

■ **12-18** Un estudio elaborado por el Departamento de Transporte de Atlanta, Georgia, acerca del efecto de los precios de boletos de autobús sobre el número de pasajeros produjo los siguientes resultados:

Precio del boleto (centavos)	25	30	35	40	45	50	55	60
Pasajeros por 100 millas	800	780	780	660	640	600	620	620

a) Grafique estos datos.
b) Desarrolle la ecuación de estimación que mejor describa estos datos.

c) Pronostique el número de pasajeros/100 millas si el precio del boleto fuera de 50 centavos. Utilice un intervalo de predicción del 95% de aproximación.

■ **12-19** William C. Andrews, consultor de comportamiento organizacional de Victory Motorcycles, ha diseñado una prueba para mostrar a los supervisores de la compañía los peligros de sobrevigilar a sus trabajadores. Un trabajador de la línea de ensamble tiene a su cargo una serie de tareas complicadas. Durante el desempeño del trabajador, un inspector lo interrumpe constantemente para ayudarlo a terminar las tareas. El trabajador, después de terminar su trabajo, recibe una prueba sicológica diseñada para medir la hostilidad del trabajador hacia la autoridad (una alta puntuación implica una hostilidad baja). A ocho distintos trabajadores se les asignaron las tareas y luego se les interrumpió para darles instrucciones útiles un número variable de veces (línea *X*). Sus calificaciones en la prueba de hostilidad se dan en el renglón *Y*.

X **(número interrupciones al trabajador)**	5	10	10	15	15	20	20	25
Y **(calificación del trabajador en la prueba de hostilidad)**	58	41	45	27	26	12	16	3

a) Grafique estos datos.
b) Desarrolle la ecuación que mejor describa la relación entre el número de interrupciones y la calificación de la prueba.
c) Pronostique la calificación esperada de la prueba si el trabajador es interrumpido 18 veces.

■ **12-20** El editor en jefe de un importante periódico metropolitano ha intentado convencer al dueño para que mejore las condiciones de trabajo en la imprenta. Está convencido de que, cuando trabajan las prensas, el grado de ruido crea niveles no saludables de tensión y ansiedad. Recientemente hizo que un sicólogo realizara una prueba durante la cual situaron a los prensistas en cuartos con niveles variables de ruido y luego les hicieron otra prueba para medir niveles de humor y ansiedad. La siguiente tabla muestra el índice de su grado de ansiedad o nerviosismo y el nivel de ruido al que se vieron expuestos (1.0 es bajo y 10.0 es alto).

Nivel de ruido	4	3	1	2	6	7	2	3
Grado de ansiedad	39	38	16	18	41	45	25	38

a) Grafique estos datos.
b) Desarrolle una ecuación de estimación que describa los datos.
c) Pronostique el grado de ansiedad que podríamos esperar cuando el nivel de ruido es 5.

■ **12-21** Una compañía administra a sus vendedores en capacitación una prueba de ventas antes de salir a trabajar. La administración de la compañía está interesada en determinar la relación entre las calificaciones de la prueba y las ventas logradas por esos vendedores al final de un año de trabajo. Se recolectaron los siguientes datos de 10 agentes de ventas que han estado en el campo un año.

Núm. de vendedor	Calif. de la prueba (*T*)	Núm. de unidades vendidas (*S*)
1	2.6	95
2	3.7	140
3	2.4	85
4	4.5	180
5	2.6	100
6	5.0	195
7	2.8	115
8	3.0	136
9	4.0	175
10	3.4	150

a) Encuentre la recta de regresión de mínimos cuadrados que podría usarse para predecir las ventas a partir de las calificaciones en la prueba de capacitación.
b) ¿En cuánto se incrementa el número esperado de unidades vendidas por cada incremento de 1 punto en una calificación de la prueba?
c) Utilice la recta de regresión de mínimos cuadrados para predecir el número de unidades que vendería un capacitando que obtuvo una calificación promedio en la prueba.

■ **12-22** El consejo municipal de la ciudad de Bowie, Maryland, ha recabado datos del número de accidentes menores de tráfico y el número de partidos de fútbol de jóvenes que tienen lugar en la ciudad el fin de semana.

X **(partidos de fútbol)**	20	30	10	12	15	25	34
Y **(accidentes menores)**	6	9	4	5	7	8	9

a) Grafique estos datos.
b) Desarrolle la ecuación de estimación que mejor describa estos datos.
c) Pronostique el número de accidentes menores de tráfico que ocurrirán en un fin de semana durante el cual tendrán lugar 33 partidos de fútbol en Bowie.
d) Calcule el error estándar de la estimación.

■ **12-23** En economía, la función de demanda de un producto a menudo se estima mediante una regresión de la cantidad vendida (Q) sobre el precio (P). La compañía Bamsy está tratando de estimar la función de demanda para su nueva muñeca "Ma'am", y ha recabado los siguientes datos:

P	20.0	17.5	16.0	14.0	12.05	10.0	8.0	6.5
Q	125	156	183	190	212	238	250	276

a) Grafique estos datos.
b) Calcule la recta de regresión de mínimos cuadrados.
c) Trace la recta de regresión ajustada en su gráfica del inciso a).

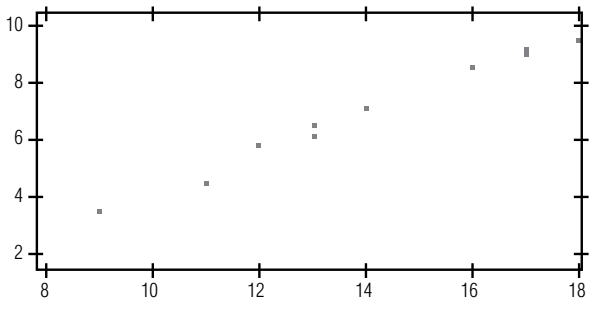
■ **12-24** Una compañía fabricante de llantas está interesada en eliminar contaminantes de los tubos de emisión de su fábrica y el costo es una preocupación. La compañía ha recolectado datos de otras compañías respecto al monto gastado en medidas ambientales y la cantidad de contaminantes eliminada que resultó (como porcentaje de la emisión total).

Dinero gastado (miles de dólares)	8.4	10.2	16.5	21.7	9.4	8.3	11.5
Porcentaje de contaminantes	35.9	31.8	24.7	25.2	36.8	35.8	33.4
Dinero gastado (miles de dólares)	18.4	16.7	19.3	28.4	4.7	12.3	
Porcentaje de contaminantes	25.4	31.4	27.4	15.8	31.5	28.9	

a) Calcule la ecuación de regresión.
b) Pronostique el porcentaje de contaminantes eliminados si se gastan $20,000 en medidas de control.
c) Calcule el error estándar de la estimación.

Soluciones a los ejercicios de autoevaluación

EA 12-2 a)



b)

X	Y	XY	X^2
13	6.2	80.6	169
16	8.6	137.6	256
14	7.2	100.8	196
11	4.5	49.5	121
17	9.0	153.0	289
9	3.5	31.5	81
13	6.5	84.5	169
17	9.3	158.1	289
18	9.5	171.0	324
12	5.7	68.4	144
$\Sigma X = 140$	$\Sigma Y = 70.0$	$\Sigma XY = 1{,}035.0$	$\Sigma X^2 = 2{,}038$

$$\overline{X} = 140/10 = 14 \qquad \overline{Y} = 70.0/10 = 7.0$$

$$b = \frac{\Sigma XY - n\overline{X}\,\overline{Y}}{\Sigma X^2 - n\overline{X}^2} = \frac{1{,}035.0 - 10(14)(7.0)}{2{,}038 - 10(14)^2} = 0.7051$$

$$a = \overline{Y} - b\overline{X} = 7.0 - (0.7051)(14) = -2.8714$$

Entonces, $\hat{Y} = -2.8714 + 0.7051X$. Si usa un paquete de regresión de computadora para hacer los cálculos, es posible que obtenga

$$\hat{Y} = -2.8718 + 0.7051X$$

Esta pequeña diferencia ocurre porque la mayoría de los paquetes de software hacen sus cálculos con más de diez lugares decimales, y aquí se redondeó b antes de calcular a. Para casi todas las situaciones prácticas, esta pequeña diferencia (es decir, $a = -2.8724$ en lugar de -2.8718) es intrascendente.

c) $X = 10$, $\hat{Y} = -2.8714 + 0.7051(10) = 4.1796$

$X = 15$, $\hat{Y} = -2.8714 + 0.7051(15) = 7.7051$

$X = 20$, $\hat{Y} = -2.8714 + 0.7051(20) = 11.2306$

EA 12-3 En este problema, Y = gastos generales y X = unidades producidas.

a)

X	Y	XY	X^2	Y^2
40	191	7,640	1,600	36,481
42	170	7,140	1,764	28,900
53	272	14,416	2,809	73,984
35	155	5,425	1,225	24,025
56	280	15,680	3,136	78,400
39	173	6,747	1,521	29,929
48	234	11,232	2,304	54,756
30	116	3,480	900	13,456
37	153	5,661	1,369	23,409
40	178	7,120	1,600	31,684
$\Sigma X = 420$	$\Sigma X 1{,}922$	$\Sigma XY = 84{,}541$	$\Sigma X^2 = 18{,}228$	$\Sigma Y^2 = 395{,}024$

$$\overline{X} = \frac{420}{10} = 42 \qquad \overline{Y} = \frac{1{,}922}{10} = 192.2$$

$$b = \frac{\Sigma XY - n\overline{X}\,\overline{Y}}{\Sigma X^2 - n\overline{X}^2} = \frac{84{,}541 - 10(42)(192.2)}{18{,}228 - 10(42)^2} = 6.4915$$

$$a = \overline{Y} - b\overline{X} = 192.2 - 6.4915(42) = -80.4430$$

Entonces, $\hat{Y} = -80.4430 + 6.4915X$ (con software: $\hat{Y} = -80.4428 + 6.4915X$).

b) $\hat{Y} = -80.4430 + 6.4915(50) = 244.1320$

c) $$s_e = \sqrt{\frac{\Sigma Y^2 - a\Sigma Y - b\Sigma XY}{n-2}}$$

$$= \sqrt{\frac{395{,}024 - (-80.4430)(1{,}922) - 6.4915(84{,}541)}{8}} = 10.2320$$

12.3 Análisis de correlación

Qué hace el análisis de correlación

El análisis de correlación es la herramienta estadística que podemos usar para describir *el grado en el que una variable está linealmente relacionada con otra.* Con frecuencia, el análisis de correlación se utiliza junto con el de regresión para medir qué tan bien la línea de regresión explica los cambios de la variable dependiente, Y. Sin embargo, la correlación también se puede usar sola para medir el grado de asociación entre dos variables.

Dos medidas que describen la correlación

Los estadísticos han desarrollado dos medidas para describir la correlación entre dos variables: el *coeficiente de determinación* y el *coeficiente de correlación*. Presentar estas dos medidas de asociación es el objetivo de esta sección.

El coeficiente de determinación

Desarrollo del coeficiente de determinación muestral

El coeficiente de determinación es la principal forma en que podemos medir el grado, o fuerza, de la asociación que existe entre dos variables, X y Y. Debido a que usamos una muestra de puntos para desarrollar rectas de regresión, nos referimos a esta medida como el *coeficiente de determinación muestral*.

El coeficiente de determinación muestral se deriva de la relación entre dos tipos de variación: la variación de los valores Y en un conjunto de datos alrededor de

1. la recta de regresión ajustada;
2. su propia media.

El término *variación* en estos dos casos se utiliza en su sentido estadístico usual para expresar "la suma de los cuadrados de un grupo de desviaciones". Usando esta definición, entonces, es razonable expresar la variación de los valores Y alrededor de la recta de regresión con esta ecuación:

Variación de los valores de *Y* alrededor de la recta de regresión

$$\text{Variación de los valores de } Y \text{ alrededor de la recta de regresión} = \Sigma(Y - \hat{Y})^2 \qquad [12\text{-}8]$$

La segunda variación, la de los valores de Y alrededor de su propia media, está determinada por:

Variación de los valores de *Y* alrededor de su propia media

$$\text{Variación de los valores de } Y \text{ alrededor de su propia media} = \Sigma(Y - \overline{Y})^2 \qquad [12\text{-}9]$$

Uno menos la razón entre estas dos variaciones es el coeficiente de determinación muestral, que se denota por r^2:

Coeficiente de determinación muestral

$$r^2 = 1 - \frac{\Sigma(Y - \hat{Y})^2}{\Sigma(Y - \overline{Y})^2} \qquad [12\text{-}10]$$

Las siguientes dos secciones mostrarán que r^2, según la definición de la ecuación 12-10, es una medida del grado de asociación lineal entre X y Y.

Una interpretación intuitiva de r^2

Considere las dos formas extremas en las que las variables X y Y pueden relacionarse. En la tabla 12-13, cada valor observado de Y cae en la línea de estimación, como puede verse en la figura 12-13. Ésta es una *correlación perfecta*.

Ecuación de estimación apropiada para un ejemplo de correlación perfecta

La ecuación de estimación apropiada para estos datos es fácil de determinar. Dado que la recta de regresión pasa por el origen, sabemos que la ordenada Y es cero; como Y se incrementa en 4 cada vez que X se incrementa en 1, la pendiente debe ser igual a 4. Por tanto, la recta de regresión es:

$$\hat{Y} = 4X$$

Tabla 12-13 Ilustración de una correlación perfecta entre dos variables, *X* y *Y*

Punto de datos	Valor de *X*	Valor de *Y*	
1st	1	4	
2nd	2	8	
3rd	3	12	
4th	4	16	
5th	5	20	
6th	6	24	$\bar{Y} = \frac{144}{8} = 18 \leftarrow$ Media de los valores de *Y*
7th	7	28	
8th	8	32	
		$\Sigma Y = \mathbf{144}$	

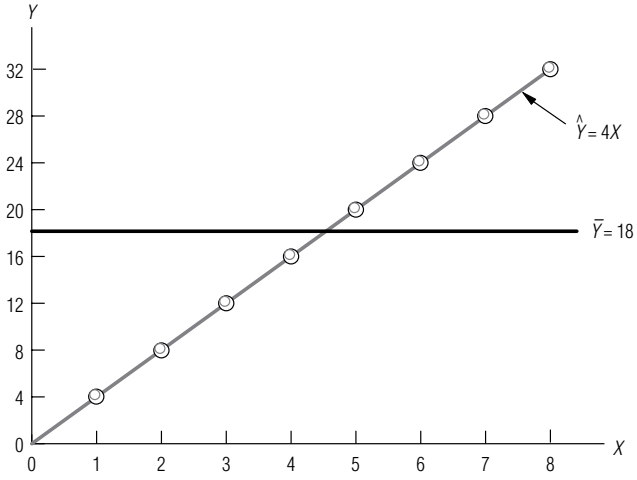


FIGURA 12-13

Correlación perfecta entre *X* y *Y*: todos los puntos caen en la recta de regresión

Desarrollo del coeficiente de determinación de la muestra para el ejemplo de una correlación perfecta

Ahora, para establecer el coeficiente de determinación de la muestra para la recta de regresión de la figura 12-13, primero calculamos el numerador de la fracción en la ecuación 12-10:

$$\text{Variación de los valores de } Y \text{ alrededor de la recta de regresión} = \Sigma(Y - \hat{Y})^2 \quad [12\text{-}8]$$
$$= \Sigma(0)^2$$
$$= 0$$

Como cada valor de *Y* está sobre la recta de regresión, la diferencia entre *Y* y $\hat{Y}$ es cero en cada caso

Entonces podemos encontrar el denominador de la fracción:

$$\text{Variación de los valores de } Y \text{ alrededor de su propia media} = \Sigma(Y - \bar{Y})^2 \quad [12\text{-}9]$$

$$(4 - 18)^2 = (-14)^2 = 196$$
$$(8 - 18)^2 = (-10)^2 = 100$$
$$(12 - 18)^2 = (-6)^2 = 36$$
$$(16 - 18)^2 = (-2)^2 = 4$$
$$(20 - 18)^2 = (2)^2 = 4$$
$$(24 - 18)^2 = (6)^2 = 36$$
$$(28 - 18)^2 = (10)^2 = 100$$
$$(32 - 18)^2 = (14)^2 = 196$$
$$\mathbf{672} \leftarrow \Sigma(Y - \bar{Y})^2$$

Tabla 12-14

Ilustración de la correlación cero entre dos variables, *X* y *Y*

Dato puntual	Valor de *X*	Valor de *Y*
1°	1	6
2°	1	12
3°	3	6
4°	3	12
5°	5	6
6°	5	12
7°	7	6
8°	7	12
		$\Sigma Y = \mathbf{72}$
		$\bar{Y} = \frac{72}{8}$
		$= 9 \leftarrow$ Media de los valores de *Y*

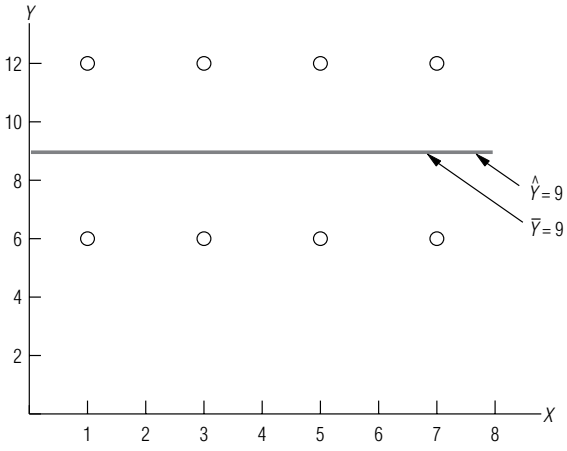


FIGURA 12-14

Correlación cero entre *X* y *Y*: los mismos valores de *Y* aparecen para distintos valores de *X*

Al sustituir estos valores en la ecuación 12-10, podemos encontrar que el coeficiente de determinación de la muestra es igual a +1:

$$r^2 = 1 - \frac{\Sigma(Y - \hat{Y})^2}{\Sigma(Y - \bar{Y})^2} \qquad [12\text{-}10]$$

$$= 1 - \frac{0}{672}$$

$$= 1 - 0$$

$$= 1 \leftarrow \text{Coeficiente de determinación de la muestra cuando hay una correlación perfecta}$$

De hecho, r^2 es igual a +l siempre que la recta de regresión sea un estimador perfecto.

Una segunda forma extrema en que las variables *X* y *Y* pueden relacionarse es aquella donde los puntos podrían caer a distancias iguales en ambos lados de una línea de regresión horizontal, como se ve en la figura 12-14. Este conjunto de datos consiste en los ocho puntos registrados en la tabla 12-14.

En la figura 12-14, podemos ver que la recta de regresión de mínimos cuadrados apropiada para estos datos está dada por la ecuación $\hat{Y} = 9$. La pendiente de la recta es *cero*, porque los mismos valores de *Y* aparecen para todos los valores de *X*. Tanto la ordenada *Y* como la media de los valores de *Y* son iguales a 9.

Cálculo del coeficiente de determinación de la muestra para la correlación cero

Ahora calcularemos las dos variaciones usando las ecuaciones 12-8 y 12-9, para poder calcular el coeficiente de determinación de la muestra para esta recta de regresión. Primero calculamos la variación de los valores de *Y* alrededor de la línea de estimación $\hat{Y} = 9$:

Variación de los valores de Y alrededor de la recta de regresión $= \Sigma(Y - \hat{Y})^2$ [12-8]

$$
\begin{aligned}
(\ 6 - 9)^2 &= (-3)^2 = 9 \\
(12 - 9)^2 &= (\ \ 3)^2 = 9 \\
(\ 6 - 9)^2 &= (-3)^2 = 9 \\
(12 - 9)^2 &= (\ \ 3)^2 = 9 \\
(\ 6 - 9)^2 &= (-3)^2 = 9 \\
(12 - 9)^2 &= (\ \ 3)^2 = 9 \\
(\ 6 - 9)^2 &= (-3)^2 = 9 \\
(12 - 9)^2 &= (\ \ 3)^2 = 9 \\
&\overline{\mathbf{72}} \leftarrow \Sigma(Y - \hat{Y})^2
\end{aligned}
$$

Variación de los valores de Y alrededor de su propia media $= \Sigma(Y - \hat{Y})^2$ [12-9]

$$
\begin{aligned}
(\ 6 - 9)^2 &= (-3)^2 = 9 \\
(12 - 9)^2 &= (\ \ 3)^2 = 9 \\
(\ 6 - 9)^2 &= (-3)^2 = 9 \\
(12 - 9)^2 &= (\ \ 3)^2 = 9 \\
(\ 6 - 9)^2 &= (-3)^2 = 9 \\
(12 - 9)^2 &= (\ \ 3)^2 = 9 \\
(\ 6 - 9)^2 &= (-3)^2 = 9 \\
(12 - 9)^2 &= (\ \ 3)^2 = 9 \\
&\overline{\mathbf{72}} \leftarrow \Sigma(Y - \bar{Y})^2
\end{aligned}
$$

Sustituyendo estos dos valores en la ecuación 12-10, vemos que el coeficiente de determinación de la muestra es 0:

$$
\begin{aligned}
r^2 &= 1 - \frac{\Sigma(Y - \hat{Y})^2}{\Sigma(Y - \bar{Y})^2} \qquad [12\text{-}10] \\
&= 1 - \frac{72}{72} \\
&= 1 - 1 \\
&= 0 \leftarrow \text{coeficiente de determinación de la muestra cuando no hay correlación}
\end{aligned}
$$

Por tanto, el valor de r^2 es cero cuando no hay correlación.

Interpretación de los valores r^2

En los problemas con que se topa la mayoría de los responsables de la toma de decisiones, r^2 caerá en alguna parte entre estos dos extremos de 1 y 0. Recuerde, no obstante, que r^2 cercana a 1 indica una fuerte correlación entre X y Y, mientras que r^2 cercana a 0 significa que existe poca correlación entre estas dos variables.

Un punto que debemos resaltar es que r^2 mide sólo la fuerza de una relación lineal entre dos variables. Por ejemplo, si tuviéramos muchos puntos X y Y, y todos cayeran en la circunferencia de un círculo, aunque dispersos aleatoriamente, claramente habría una relación entre estos puntos (todos están en el mismo círculo). Pero en este caso, si calculáramos r^2, resultaría estar cerca de cero, porque los puntos no tienen una relación *lineal* entre ellos.

Otra interpretación de r^2

Otra forma de interpretar el coeficiente de determinación de la muestra

Los estadísticos también interpretan el coeficiente de determinación de la muestra viendo la *cantidad de la variación en Y que se explica por la recta de regresión.* Para entender este significado de r^2, consideremos la recta de regresión de la figura 12-15. Aquí, separamos un valor observado de Y, mostrado como el círculo negro superior. Si usamos la media de los valores de Y, $\bar{Y}$, para estimar este

Desviación explicada y no explicada

valor de *Y*, del círculo negro, entonces la *desviación total* de esta *Y* lejos de su media sería $(Y - \overline{Y})$. Observe que si usamos la recta de regresión para estimar este valor de *Y* del círculo negro, obtendríamos una mejor estimación. Sin embargo, aun cuando la recta de regresión justifica, o explica $(\hat{Y} - Y)$ de la desviación total, la porción restante de la desviación total $(Y - \hat{Y})$ sigue *sin explicarse*.

Pero consideremos un conjunto completo de valores *Y* observados en vez de un solo valor. La variación total, esto es, la suma de los cuadrados de las desviaciones totales, de estos puntos alrededor de su media sería:

$$\Sigma(Y - \overline{Y})^2 \qquad [12\text{-}9]$$

Variación explicada y no explicada

y la porción *explicada* de la variación total, o la suma de los cuadrados de las desviaciones explicadas de estos puntos alrededor de su media, sería:

$$\Sigma(\hat{Y} - \overline{Y})^2$$

La porción *no explicada* de la variación total (la suma de los cuadrados de las desviaciones no explicadas) de estos puntos respecto a su recta de regresión sería:

$$\Sigma(Y - \hat{Y})^2 \qquad [12\text{-}8]$$

Si deseamos expresar la fracción de la variación total que queda *no explicada*, dividiríamos la variación no explicada, $\Sigma(Y - \hat{Y})^2$, entre la variación total, $\Sigma(Y - \overline{Y})^2$, de la siguiente manera:

$$\frac{\Sigma(Y - \hat{Y})^2}{\Sigma(Y - \overline{Y})^2} \leftarrow \text{Fracción de la variación total no explicada}$$

y**,** finalmente, si restamos de 1 la fracción de la variación total que sigue no explicada, tendremos la fórmula para encontrar la fracción de la variación total de *Y* que *es* explicada por la recta de regresión. Esa fórmula es:

$$r^2 = 1 - \frac{\Sigma(Y - \hat{Y})^2}{\Sigma(Y - \overline{Y})^2} \qquad [12\text{-}10]$$

la misma ecuación que usamos para calcular r^2. Es en este sentido que r^2 mide qué tan bien *X* explica *Y*, esto es, el grado de asociación entre *X* y *Y*.

Método abreviado para calcular r^2

Una observación final respecto al cálculo de r^2. Para obtener r^2 usando las ecuaciones 12-8, 12-9 y 12-10, se requiere una serie de cálculos tediosos; para evitarlos, los estadísticos han desarrollado una versión abreviada, usando valores que habríamos determinado de antemano en el análisis de regresión. La fórmula es:

Método abreviado para obtener el coeficiente de determinación de la muestra

$$r^2 \text{ calculada por el método corto} \longrightarrow r^2 = \frac{a\Sigma Y + b\Sigma XY - n\overline{Y}^2}{\Sigma Y^2 - n\overline{Y}^2} \qquad [12\text{-}11]$$

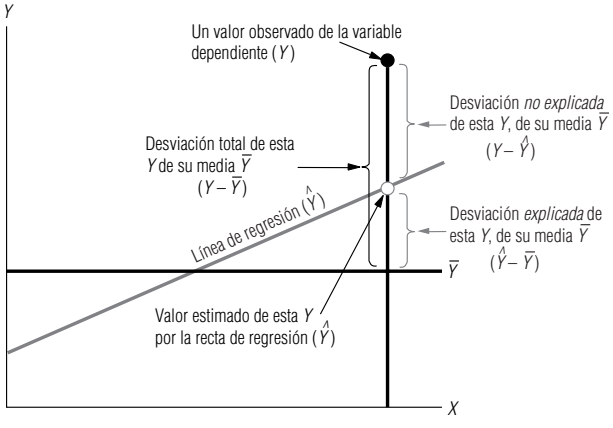


FIGURA 12-15

Desviación total, desviación explicada y desviación no explicada para un valor observado de *Y*

Tabla 12-15

Cálculo de los datos para la ecuación 12-11

Año ($n = 6$) (1)	Gastos de ID (X) (2)	Ganacia anual (Y) (3)	XY (2) × (3)	X^2 $(2)^2$	Y^2 $(3)^2$
1995	5	31	155	25	961
1994	11	40	440	121	1,600
1993	4	30	120	16	900
1992	5	34	170	25	1,156
1991	3	25	75	9	625
1990	2	20	40	4	400
	$\Sigma X = \mathbf{30}$	$\Sigma Y = \mathbf{180}$	$\Sigma XY = \mathbf{1,000}$	$\Sigma X^2 = \mathbf{200}$	$\Sigma Y^2 = \mathbf{5,642}$

$$\bar{Y} = \frac{180}{6}$$

$= 30 \leftarrow$ Media de los valores de la variable dependiente

donde,

- $r^2 =$ coeficiente de determinación de la muestra
- $a =$ ordenada Y
- $b =$ pendiente de la línea de estimación de mejor ajuste
- $n =$ número de puntos de datos
- $X =$ valores de la variable independiente
- $Y =$ valores de la variable dependiente
- $\bar{Y} =$ media de los valores observados de la variable dependiente

Aplicación del método abreviado

Para ver por qué esta fórmula constituye un método abreviado, la aplicaremos a la regresión que relaciona los gastos de investigación y desarrollo con las ganancias. En la tabla 12-15, repetimos las columnas de la tabla 12-9, añadiendo una columna Y^2. Recuerde que cuando encontramos los valores para a y b, la recta de regresión para este problema era:

$$\hat{Y} = 20 + 2X$$

Usando esta recta y la información de la tabla 12-15, podemos calcular r^2 de la siguiente manera:

$$r^2 = \frac{a\Sigma Y + b\Sigma XY - n\bar{Y}^2}{\Sigma Y^2 - n\bar{Y}^2} \qquad [12\text{-}11]$$

$$= \frac{(20)(180) + (2)(1{,}000) - (6)(30)^2}{5{,}642 - (6)(30)^2}$$

$$= \frac{3{,}600 + 2{,}000 - 5{,}400}{5{,}642 - 5{,}400}$$

$$= \frac{200}{242}$$

$= 0.826 \leftarrow$ Coeficiente de determinación de la muestra

Interpretación de r^2

Así, podemos concluir que la variación en los gastos de investigación y desarrollo (la variable independiente X) explica el 82.6% de la variación en las ganancias anuales (la variable dependiente Y).

El coeficiente de correlación

Coeficiente de correlación de la muestra

El coeficiente de correlación es la segunda medida que podemos usar para describir qué tan bien explica una variable a otra. Cuando tratamos con muestras, el *coeficiente de correlación de la muestra* se denota por r y es la raíz cuadrada del coeficiente de determinación de muestra:

Coeficiente de correlación de la muestra

$$r = \sqrt{r^2} \qquad [12\text{-}12]$$

Cuando la pendiente de la ecuación de estimación es positiva, r es la raíz cuadrada positiva, pero si b es negativa, r es la raíz cuadrada negativa. Entonces, **el signo de *r* indica la dirección de la relación entre las dos variables *X* y *Y***. Si existe una relación inversa —esto es, si Y disminuye al aumentar X—, entonces r caerá entre 0 y -1. De manera similar, si existe una relación directa (si Y aumenta al aumentar X), entonces r será un valor en el intervalo de 0 a 1. La figura 12-16 ilustra estas características de r.

Interpretación de *r*

El coeficiente de correlación es más difícil de interpretar que r^2. ¿Qué significa $r = 0.9$? Para responder esta pregunta, debemos recordar que $r = 0.9$ es lo mismo que $r^2 = 0.81$. Esto último nos dice que el 81% de la variación en Y es explicada por la recta de regresión. De esta forma, vemos que r es sólo la raíz cuadrada de r^2, y su significado es qué tanto se relacionan las variables x y y. Por lo que $r = 0.9$ significa que el 90% de los datos se relacionan entre sí.

Cálculo de *r* para el problema de investigación y desarrollo

Ahora encontremos el coeficiente de correlación del problema que relaciona gastos de investigación y desarrollo con ganancias anuales. En la sección anterior, encontramos que el coeficiente de determinación de la muestra es $r^2 = 0.826$, de manera que podemos sustituir este valor en la ecuación 12-12 y encontrar que

$$\begin{aligned} r &= \sqrt{r^2} && [12\text{-}12] \\ &= \sqrt{0.826} \\ &= 0.909 \leftarrow \text{Coeficiente de correlación de la muestra} \end{aligned}$$

La relación entre las dos variables es directa y la pendiente es positiva; por tanto, el signo de r es positivo.

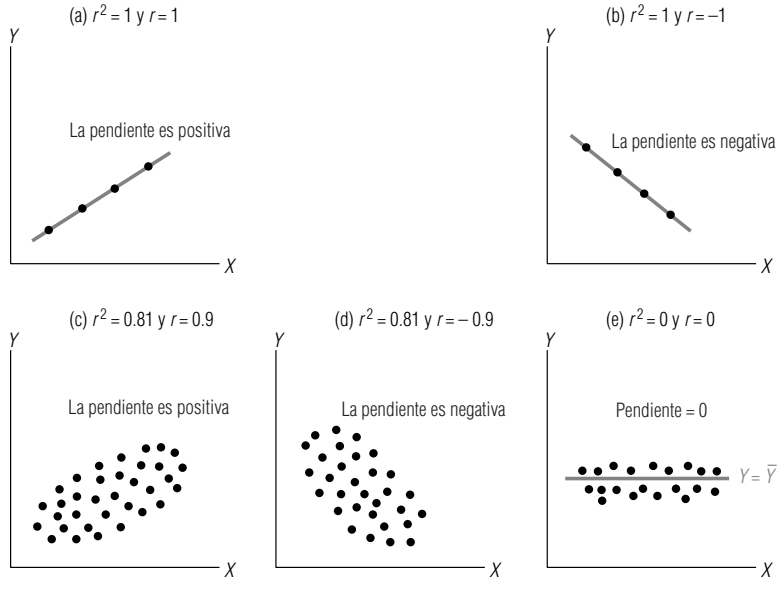


FIGURA 12-16

Varias características de *r*, el coeficiente de correlación de la muestra

SUGERENCIAS Y SUPOSICIONES Advertencia: como ya sabe que el coeficiente de determinación (r^2) es el cuadrado del coeficiente de correlación, r, debe tener cuidado de usar todo, menos las correlaciones más altas, como base para tomar decisiones. Sugerencia: si se encuentra que la cantidad gastada en películas se correlaciona a 0.6 con el ingreso familiar, parece una correlación bastante fuerte (0.6 está más cerca de 1.0 que de cero). Pero cuando se eleva al cuadrado, se ve que es responsable sólo de $0.6 \times 0.6 = 0.36$ o el 36% de la variación en la cantidad de dinero que gastan las familias en películas. Si diseña su estrategia de mercado dirigida sólo a familias con altos ingresos perderá muchos clientes potenciales. Sugerencia: en su lugar, intente averiguar qué más influye en las decisiones de películas de las familias.

Ejercicios 12.3

Ejercicios de autoevaluación

EA 12-4 Las librerías de la universidad han vendido el libro *Believe or Not: Wonders of Statistics Guide* durante 12 semestres y desean estimar la relación entre las ventas y el número de secciones de estadística elemental que se enseñan en cada semestre. Se recolectaron los siguientes datos:

Ventas (unidades)	33	38	24	61	52	45
Número de secciones	3	7	6	6	10	12
Ventas (unidades)	65	82	29	63	50	79
Número de secciones	12	13	12	13	14	15

a) Desarrolle la ecuación de estimación que mejor se ajuste a los datos.
b) Calcule el coeficiente de determinación de la muestra y el coeficiente de correlación de la muestra.

EA 12-5 Calcule el coeficiente de determinación de la muestra y el coeficiente de correlación de la muestra para los datos del ejercicio EA 12-3.

Conceptos básicos

■ **12-25** ¿Qué tipo de correlación (positiva, negativa o cero) debe esperarse de estas variaciones?
a) Habilidad de los supervisores y producción de sus subordinados.
b) Edad en el primer trabajo de tiempo completo y años de educación.
c) Peso y presión sanguínea.
d) Promedio general en la universidad y estatura del estudiante.

En los siguientes ejercicios, calcule el coeficiente de determinación de la muestra y el coeficiente de correlación para los ejercicios especificados.

■ **12-26** Calcule el coeficiente de determinación de la muestra y el coeficiente de correlación para los datos del ejercicio 12-17.

■ **12-27** Calcule el coeficiente de determinación de la muestra y el coeficiente de correlación para los datos del ejercicio 12-18.

■ **12-28** Calcule el coeficiente de determinación de la muestra y el coeficiente de correlación para los datos del ejercicio 12-19.

■ **12-29** Calcule el coeficiente de determinación de la muestra y el coeficiente de correlación para los datos del ejercicio 12-20.

■ **12-30** Calcule el coeficiente de determinación de la muestra y el coeficiente de correlación para los datos del ejercicio 12-21.

Aplicaciones

■ **12-31** El Bank of Lincoln está interesado en reducir el tiempo que las personas esperan para ver a su banquero personal. También le interesa la relación entre el tiempo de espera (Y) en minutos y el número de banqueros atendiendo (X). Los clientes se seleccionaron al azar con los datos siguientes:

X	2	3	5	4	2	6	1	3	4	3	3	2	4
Y	12.8	11.3	3.2	6.4	11.6	3.2	8.7	10.5	8.2	11.3	9.4	12.8	8.2

a) Calcule la ecuación de regresión que mejor se ajusta a estos datos.
b) Calcule el coeficiente de determinación de la muestra y el coeficiente de correlación de muestra.

■ **12-32** Zippy Cola está estudiando el efecto de su última campaña publicitaria. Se escogieron personas al azar y se les llamó para preguntarles cuántas latas de Zippy Cola habían comprado la semana anterior y cuántos anuncios de Zippy Cola habían leído o visto durante el mismo periodo.

X (número de anuncios)	3	7	4	2	0	4	1	2
Y (latas compradas)	11	18	9	4	7	6	3	8

a) Desarrolle la ecuación de estimación que mejor ajuste los datos.
b) Calcule el coeficiente de determinación de la muestra y el coeficiente de correlación.

Soluciones a los ejercicios de autoevaluación

EA 12-4 En este problema, Y = ventas y X = número de secciones.

a)

X	Y	XY	X^2	Y^2
3	33	99	9	1,089
7	38	266	49	1,444
6	24	144	36	576
6	61	366	36	3,721
10	52	520	100	2,704
12	45	540	144	2,025
12	65	780	144	4,225
13	82	1,066	169	6,724
12	29	348	144	841
13	63	819	169	3,969
14	50	700	196	2,500
15	79	1,185	225	6,241
$\Sigma X =$ **123**	$\Sigma Y =$ **621**	$\Sigma XY =$ **6,833**	$\Sigma X^2 =$ **1,421**	$\Sigma Y^2 =$ **36,059**

$$\overline{X} = 123/12 = 10.25 \qquad \overline{Y} = 621/12 = 51.75$$

$$b = \frac{\Sigma XY - n\overline{X}\,\overline{Y}}{\Sigma X^2 - n\overline{X}^2} = \frac{6{,}833 - 12(10.25)(51.75)}{1{,}421 - 12(10.25)^2} = 2.9189$$

$$a = \overline{Y} - b\overline{X} = 51.75 - 2.9189(10.25) = 21.8313$$

Entonces, $\hat{Y} = 21.8313 + 2.9189X$ (con software: $\hat{Y} = 21.8315 + 2.9189X$).

b) $$r^2 = \frac{a\Sigma Y + b\Sigma XY - n\overline{Y}^2}{\Sigma Y^2 - n\overline{Y}^2}$$

$$= \frac{21.8313(621) + 2.9189(6{,}833) - 12(51.75)^2}{36{,}059 - 12(51.75)^2} = 0.3481$$

$$r = \sqrt{0.3481} = 0.5900$$

EA 12-5 De la solución del ejercicio EA 12-3, se tiene $n = 10$, $\Sigma Y = 1{,}922$, $\overline{Y} = 192.2$, $\Sigma XY = 84{,}541$, $\Sigma Y^2 = 395{,}024$, $a = -80.4430$ y $b = 6.4915$. Por tanto,

$$r^2 = \frac{a\Sigma Y + b\Sigma XY - n\overline{Y}^2}{\Sigma Y^2 - n\overline{Y}^2}$$

$$= \frac{-80.4430(1{,}922) + 6.4915(84{,}541) - 10(192.2)^2}{395{,}024 - 10(192.2)^2}$$

$$= 0.9673$$

$$r = \sqrt{0.9673} = 0.9835$$

12.4 Inferencias sobre parámetros de población

Relación de la recta de regresión de la muestra y la recta de regresión de la población

Hasta ahora, hemos usado los análisis de regresión y correlación para relacionar dos variables con base en la información de la muestra. Pero los datos de una muestra sólo representan una parte de la población total. Debido a esto, podemos concebir nuestra recta de regresión de la muestra estimada como una estimación de una recta de regresión de la población verdadera, aunque desconocida, de la forma:

Recta de regresión de la población

$$Y = A + BX \qquad [12\text{-}13]$$

Recuerde nuestro problema acerca de la directora del Departamento de Salubridad que trataba de usar la antigüedad de un camión para explicar su gasto anual de reparaciones. Ese gasto probablemente consiste en dos partes:

1. Mantenimiento regular independiente de la antigüedad del camión: afinación, cambio de aceite y lubricación. Este gasto es captado en el término de la ordenada A de la ecuación 12-13.
2. Gastos por reparaciones debidos a la antigüedad: realineación de frenos, revisión de motor y transmisión, y pintura. Tales gastos tenderán a incrementarse con la antigüedad del camión, y son captados en el término BX de la recta de regresión de la población $Y = A + BX$ de la ecuación 12-13.

¿Por qué los datos puntuales (o puntos) no caen exactamente en la recta de regresión?

Claro está que no todos los frenos de todos los camiones se desgastan al mismo tiempo, y algunos de los camiones funcionarán durante años sin revisiones de motor. Debido a esto, los puntos individuales probablemente no caerán exactamente en la recta de regresión de población. Algunos estarán arriba; otros, abajo. Así que, en vez de satisfacer

$$Y = A + BX \qquad [12\text{-}13]$$

los puntos individuales satisfarán la fórmula:

Recta de regresión de la población con variación aleatoria

$$Y = A + BX + e \qquad [12\text{-}13a]$$

Variación aleatoria e y su comportamiento

donde e es una perturbación o variación aleatoria de la recta de regresión de la población. En promedio, e es igual a cero, porque las variaciones arriba de la recta de regresión poblacional se anulan con las variaciones abajo de esa recta. Podemos expresar la desviación estándar de estas variaciones individuales mediante σ_e. El error estándar de la estimación s_e, entonces, es una estimación de σ_e, la desviación estándar de las variaciones.

Veamos con más cuidado las ecuaciones 12-13 y 12-13a. La ecuación 12-13a expresa los valores de Y (en este caso, el gasto anual de reparaciones) en términos de los valores individuales de X (la antigüedad de un camión) y la variación aleatoria (e). Puesto que las variaciones arriba de la recta de regresión de población se anulan por aquéllas situadas abajo, sabemos que el valor esperado de e es cero, y vemos que si tuviéramos varios camiones de la misma antigüedad, X, esperaríamos que el gasto anual de reparaciones para estos camiones fuera $Y = A + BX$. Esto nos muestra que la recta de regresión de la población (ecuación 12-13) proporciona el valor medio de Y asociado con cada valor de X.

Inferencias sobre B a partir de b

Puesto que nuestra recta de regresión de la *muestra*, $\hat{Y} = a + bX$ (ecuación 12-3), estima la recta de regresión de la *población*, $Y = A + BX$ (ecuación 12-13), deberíamos poder usarla para hacer inferencias acerca de la recta de regresión de la población. Entonces, en esta sección haremos infe-

rencias respecto a la pendiente B de la ecuación de regresión "verdadera" (de toda la población), basadas en la pendiente b de la ecuación de regresión estimada a partir de una muestra de valores.

Pendiente de la recta de regresión de la población

Diferencia entre la ecuación de regresión verdadera y la estimada a partir de observaciones de la muestra

La recta de regresión se deriva de una muestra y no de una población entera. Como resultado, no podemos esperar que la ecuación de regresión, $Y = A + BX$ (de toda la población), sea exactamente la misma que la ecuación estimada a partir de observaciones de la muestra, o $\hat{Y} = a + bX$. Aun así, podemos usar el valor de b, la pendiente que calculamos a partir de una muestra para probar hipótesis respecto al valor de B, la pendiente de la recta de regresión para toda la población.

Prueba de hipótesis respecto a *B*

El procedimiento para probar una hipótesis respecto a B es similar a los procedimientos presentados en los capítulos 8 y 9, de pruebas de hipótesis. Para comprender este proceso, regresemos al problema de la relación entre los gastos anuales de investigación y desarrollo, y las ganancias. En la página 524, señalamos que $b = 2$. El primer paso es encontrar algún valor para B con el fin de compararlo con $b = 2$.

Supongamos que durante un periodo extenso, la pendiente de la relación entre X y Y fue 2.1. Para probar si éste es todavía el caso, podríamos definir las hipótesis como

$$H_0: B = 2.1 \leftarrow \text{Hipótesis nula}$$

$$H_1: B \neq 2.1 \leftarrow \text{Hipótesis alternativa}$$

Error estándar del coeficiente de regresión

Entonces, de hecho estamos probando para saber si los datos actuales indican que B ha cambiado de su valor histórico de 2.1.

Para encontrar el estadístico de prueba para B, es necesario primero encontrar el *error estándar del coeficiente de regresión*. Aquí, el coeficiente de regresión con el que estamos trabajando es b, así que el error estándar de este coeficiente se expresa como s_b. La ecuación 12-14 presenta la fórmula matemática para s_b:

Error estándar de *b*

$$s_b = \frac{s_e}{\sqrt{\Sigma X^2 - n\bar{X}^2}} \qquad [12\text{-}14]$$

donde,

- s_b = error estándar del coeficiente de regresión
- s_e = error estándar de la estimación
- X = valores de la variable independiente
- $\bar{X}$ = media de los valores de la variable independiente
- n = número de datos

Estandarización del coeficiente de regresión

Una vez calculado s_b, podemos utilizar la ecuación 12-15 para estandarizar la pendiente de nuestra ecuación de regresión:

Valor estandarizado de *b*

$$t = \frac{b - B_{H_0}}{s_b} \qquad [12\text{-}15]$$

en la que,

- b = pendiente de la regresión ajustada
- B_{H_0} = pendiente real hipotética para la población
- s_b = error estándar del coeficiente de regresión

Como la prueba estará basada en la distribución t con $n - 2$ grados de libertad, usamos t para denotar la estadística estandarizada.

Un vistazo a la tabla 12-15 nos permite calcular los valores de ΣX^2 y $n\overline{X}^2$. Para obtener s_e, podemos tomar un método abreviado, de la siguiente manera:

Cálculo de s_e

$$s_e = \sqrt{\frac{\Sigma Y^2 - a\Sigma Y - b\Sigma XY}{n-2}} \qquad [12\text{-}7]$$

$$= \sqrt{\frac{5{,}642 - (20)(180) - (2)(1{,}000)}{6-2}}$$

$$= \sqrt{\frac{42}{4}}$$

$$= 10.5$$

$$= 3.24 \leftarrow \text{Error estándar de la estimación}$$

Ahora podemos determinar el error estándar del coeficiente de regresión:

Cálculo de s_b

$$s_b = \frac{s_e}{\sqrt{\Sigma X^2 - n\overline{X}^2}} \qquad [12\text{-}14]$$

$$= \frac{3.24}{200 - (6)(5)^2}$$

$$= \frac{3.24}{\sqrt{50}}$$

$$= \frac{3.24}{7.07}$$

$$= 0.46 \leftarrow \text{Error estándar del coeficiente de regresión}$$

Estandarización del coeficiente de regresión

Ahora usamos el error estándar del coeficiente de regresión para calcular el estadístico de prueba estandarizado:

$$t = \frac{b - B_{H_0}}{s_b} \qquad [12\text{-}15]$$

$$= \frac{2.0 - 2.1}{0.46}$$

$$= -0.217 \leftarrow \text{Coeficiente de regresión estandarizado}$$

Conducción de la prueba de hipótesis

Suponga que tenemos razones para probar nuestra hipótesis al 10% de nivel de significancia. Como tenemos seis observaciones en nuestra muestra, sabemos que tenemos $n - 2$ o $6 - 2 = 4$ grados de libertad. Consultamos la tabla 2 del apéndice bajo la columna de 10% y bajamos hasta encontrar el renglón de 4 grados de libertad. Allí vemos que el valor t adecuado es 2.132. Puesto que nos interesa si b (la pendiente de la recta de regresión de la muestra) es significativamente *diferente* de B (la pendiente hipotética de la recta de regresión de la población), ésta es una prueba de dos colas, y los valores críticos son ± 2.132. El coeficiente de regresión estandarizado es -0.217, que está *dentro* de la región de aceptación de nuestra prueba de hipótesis. Por tanto, aceptamos la hipótesis nula de que B sigue siendo igual a 2.1. En otras palabras, no existe suficiente diferencia entre b y 2.1 para que concluyamos que B ha cambiado de su valor histórico. Por esto, sentimos que cada millón de dólares adicional gastado en investigación y desarrollo todavía aumentará las ganancias anuales aproximadamente \$2.1 millones, como sucedía en el pasado.

Además de la prueba de hipótesis, también podemos construir un *intervalo de confianza* para el valor de *B*. De la misma forma que *b* es una estimación puntual de *B*, estos intervalos de confianza son estimaciones de intervalo de *B*. El problema que acabamos de resolver, y para el cual hicimos una prueba de hipótesis, ilustrará el proceso de construir un intervalo de confianza. Encontramos que:

$$b = 2.0$$
$$s_b = 0.46$$
$$t = 2.132 \leftarrow \text{Nivel de significancia del l0\% y 4 grados de libertad}$$

Intervalo de confianza para *B*

Con esta información, podemos calcular intervalos de confianza como éste:

$$\begin{aligned} b + t(s_b) &= 2 + (2.132)(0.46) \\ &= 2 + 0.981 \\ &= 2.981 \leftarrow \text{Límite superior} \end{aligned}$$

$$\begin{aligned} b - t(s_b) &= 2 - (2.132)(0.46) \\ &= 2 - 0.981 \\ &= 1.019 \leftarrow \text{Límite inferior} \end{aligned}$$

Interpretación del intervalo de confianza

En esta situación, entonces, estamos 90% seguros de que el valor verdadero de *B* cae entre 1.019 y 2.981, esto es, cada millón de dólares adicional gastado en investigación y desarrollo incrementa las ganancias anuales en una cantidad entre \$1.02 millones y \$2.98 millones.

SUGERENCIAS Y SUPOSICIONES

En esta sección se usaron observaciones de la muestra para calcular *b*, la pendiente de la recta de regresión de la *muestra*, que después utilizamos para probar la hipótesis acerca de *B*, la pendiente verdadera de la recta de regresión de la *población*. Sugerencia: se usa s_e para calcular el error estándar del coeficiente de regresión tal como se usó la desviación estándar para calcular el error estándar de la media en el capítulo 6. Advertencia: cuando use su computadora para desarrollar una recta de regresión, no olvide preguntar, "¿es este coeficiente de regresión significativamente diferente de cero?". Si *no* lo es, no importa qué tan bien se vea la salida de la computadora, no ha demostrado una relación significativa entre las variables y deberá seguir buscando relaciones más útiles. Por ejemplo, si tiene un salón para bronceado con luz ultravioleta y tiene idea de que llegan más personas en días nublados, puede hacer una regresión del "número de visitas" con las "horas de sol". Si lo hace y obtiene una recta de regresión con una pendiente que *no* es significativa, estar al tanto del clima no le ayudará en su negocio.

Ejercicios 12.4

Ejercicios de autoevaluación

EA 12-6 En finanzas, es de interés observar la relación entre *Y*, el rendimiento promedio de las acciones, y *X*, el rendimiento global del mercado. El coeficiente de la pendiente calculada por una regresión lineal se conoce como la *beta* de las acciones por los analistas de inversiones. Una beta mayor que 1 indica que la acción es relativamente sensible a cambios en el mercado, mientras que una beta menor que 1 indica que la acción es relativamente insensible. Para los datos siguientes, calcule la beta y pruebe si ésta es significativamente menor que 1. Use $\alpha = 0.05$.

Y (%)	10	12	8	15	9	11	8	10	13	11
X (%)	11	15	3	18	10	12	6	7	18	13

EA 12-7 En un problema de regresión con un tamaño de muestra de 17, se encontró que la pendiente era 3.73 y el error estándar de la estimación era 28.654. La cantidad $(\Sigma X^2 - n\overline{X}^2) = 871.56$.

a) Encuentre el error estándar del coeficiente de la pendiente de regresión.

b) Construya un intervalo de confianza del 98% para la pendiente de la población.
c) Interprete el intervalo de confianza de la parte b).

Conceptos básicos

■ **12-33** En un problema de regresión con un tamaño de muestra de 25, se encontró que la pendiente es 1.12 y el error estándar de la estimación, 8.516. La cantidad $(\Sigma X^2 - n\bar{X}^2) = 327.52$.
a) Encuentre el error estándar del coeficiente de pendiente de regresión.
b) Pruebe si el coeficiente de regresión es diferente de 0 para un nivel de significancia de 0.05.
c) Construya un intervalo de confianza del 95% para la pendiente de la población.

Aplicaciones

■ **12-34** Ned's Beds está considerando contratar a una compañía de publicidad para estimular el negocio. Fred, el hermano de Ned, investigó el campo de la publicidad de camas y recolectó los siguientes datos de la cantidad de ganancias (Y) que logra una compañía de camas y la cantidad gastada en publicidad (X). Si Fred calcula la ecuación de regresión, la pendiente de la recta indicará el incremento en la ganancia por dólar gastado en publicidad. Ned hará la publicidad sólo si la ganancia de cada $1 invertido excede $1.50. Calcule la pendiente de la ecuación de regresión y pruebe si es mayor que 1.50. Para un nivel de significancia de 0.05, ¿debe Ned hacer la publicidad?

Cantidad de publicidad (X), en cientos de dólares	3.6	4.8	9.7	12.6	11.5	10.9
Ganancia (Y), en cientos de dólares	12.13	14.7	22.83	28.4	28.33	27.05
Cantidad de publicidad (X), en cientos de dólares	14.6	18.2	3.7	9.8	12.4	16.9
Ganancia (Y), en cientos de dólares	33.6	40.8	9.4	24.84	30.17	34.7

■ **12-35** Un corredor de una empresa de inversión local ha estudiado la relación entre el incremento en el precio del oro (X) y las peticiones de sus clientes de liquidar las acciones (Y). Del conjunto de datos basado en 15 observaciones, se encontró que la pendiente de la muestra era 2.9. Si el error estándar del coeficiente de la pendiente de regresión es 0.18, ¿existe una razón para pensar (a 0.05 de nivel de significancia) que la pendiente cambió de su valor anterior de 3.2?

■ **12-36** Para una muestra de 25, se encontró que la pendiente era 1.685 y el error estándar del coeficiente de regresión era 0.11. ¿Hay razones para creer que la pendiente ha cambiado de su valor anterior de 1.50? Utilice el nivel de significancia de 0.05.

■ **12-37** Los corredores de bienes raíces a menudo están interesados en ver cómo el avalúo de una casa varía de acuerdo con su tamaño. A continuación se muestran algunos datos del área (en miles de pies cuadrados) y el avalúo (en miles de dólares) para una muestra de 11 casas.

Área	1.1	1.5	1.6	1.6	1.4	1.3	1.1	1.7	1.9	1.5	1.3
Valor	75	95	110	102	95	87	82	115	122	98	90

a) Estime la regresión de mínimos cuadrados para predecir el valor según el avalúo a partir del tamaño
b) Generalmente, los corredores de bienes raíces sienten que el valor de una casa sube 50,000 dólares por cada 1,000 pies cuadrados de área. Para esta muestra, ¿se cumple esta relación? Utilice $\alpha = 0.10$.

■ **12-38** En 1969, una agencia de salud del gobierno estadounidense encontró que en cierto número de condados, la relación entre fumadores y muertes, por enfermedades del corazón, por cada 100,000 habitantes tenía una pendiente de 0.08. Un estudio reciente de 18 condados produjo una pendiente de 0.147 y un error estándar del coeficiente de pendiente de regresión de 0.032.
a) Construya una estimación del intervalo de confianza del 90% para la pendiente de la recta de regresión verdadera. ¿El resultado de este estudio indica que la pendiente verdadera ha cambiado?

b) Construya una estimación de intervalo de confianza del 99% para la pendiente de la recta de regresión verdadera. ¿Indica el resultado de este estudio que la pendiente verdadera ha cambiado?

■ **12-39** La compañía local de teléfonos siempre ha supuesto que el número promedio de llamadas diarias aumenta en 1.5 por cada persona adicional en una casa. Se ha sugerido que la gente es más platicadora que lo que esto refleja. Se tomó una muestra de 64 casas y se calculó que la pendiente de regresión de Y (número promedio de llamadas diarias) sobre X (tamaño de la casa) era 1.8 con un error estándar del coeficiente de la pendiente de regresión de 0.2. Pruebe si se hacen significativamente más llamadas por persona adicional de lo que la compañía de teléfonos supone; use $\alpha = 0.05$. Establezca las hipótesis y la conclusión explícitas.

■ **12-40** Los funcionarios universitarios responsables de la admisión constantemente buscan variables con las cuales predecir los promedios de calificaciones de los aspirantes. Una variable de uso común es el promedio de calificaciones del bachillerato. Para una universidad, los datos anteriores indicaban que la pendiente era 0.85. Un pequeño estudio reciente de 20 estudiantes encontró que la pendiente de la muestra era 0.70 y que el error estándar de la estimación era 0.60. La cantidad $(\Sigma X^2 - n\overline{X}^2)$ era igual que 0.25. Al nivel de significancia de 0.01, ¿debería concluir la universidad que la pendiente ha cambiado?

Soluciones a los ejercicios de autoevaluación

EA 12-6

X	Y	XY	X^2	Y^2
11	10	110	121	100
15	12	180	225	144
3	8	24	9	64
18	15	270	324	225
10	9	90	100	81
12	11	132	144	121
6	8	48	36	64
7	10	70	49	100
18	13	234	324	169
13	11	143	169	121
$\Sigma X = \mathbf{113}$	$\Sigma Y = \mathbf{107}$	$\Sigma XY = \mathbf{1{,}301}$	$\Sigma X^2 = \mathbf{1{,}501}$	$\Sigma Y^2 = \mathbf{1{,}189}$

$$\overline{X} = \frac{113}{10} = 11.3 \qquad \overline{Y} = \frac{107}{10} = 10.7$$

$$b = \frac{\Sigma XY - n\overline{X}\,\overline{Y}}{\Sigma X^2 - n\overline{X}^2} = \frac{1{,}301 - 10(11.3)(10.7)}{1{,}501 - 10(11.3)^2} = 0.4101$$

$$a = \overline{Y} - b\overline{X} = 10.7 - 0.4101(11.3) = 6.0659$$

(con software: 6.0660)

$$s_e = \sqrt{\frac{\Sigma Y^2 - a\Sigma Y - b\Sigma XY}{n-2}}$$

$$= \sqrt{\frac{1{,}189 - 6.0659(107) - 0.4101(1{,}301)}{8}} = 0.8950$$

(con software: 0.8953)

$$s_b = \frac{s_e}{\sqrt{\Sigma X^2 - n\overline{X}^2}} = \frac{0.8950}{\sqrt{224.1}} = 0.060$$

$H_0: B = 1 \qquad H_1: B < 1 \qquad \alpha = 0.05$

El estadístico estandarizado es $t = \dfrac{b - B_{H_0}}{s_b} = \dfrac{0.4101 - 1}{0.06} = -9.83$. Debido a que el valor crítico de $t(-1.860)$ es mayor que -9.83, se rechaza H_0. Las acciones son insensibles a los cambios en el mercado (la pendiente es significativamente < 1).

EA 12-7 a) $s_b = \dfrac{s_e}{\sqrt{\Sigma X^2 - n\overline{X}^2}} = \dfrac{28.654}{\sqrt{871.56}} = 0.9706$

b) El intervalo de confianza del 98% es

$$b \pm t(s_b) = 3.73 \pm 2.602(0.9706) = 3.73 \pm 2.53 = (1.20, 6.26).$$

c) En el muestreo repetido, 98 de cada 100 intervalos construidos como se acaba de hacer contienen la pendiente verdadera desconocida de la población, *B*. Para una sola muestra, se puede decir que se tiene el 98% de seguridad de que el intervalo calculado contiene a *B*.

12.5 Uso del análisis de regresión y correlación: limitaciones, errores y advertencias

Uso incorrecto de regresión y correlación

Los análisis de regresión y correlación son herramientas estadísticas que, cuando se utilizan de forma correcta, pueden prestar una ayuda significativa a las personas que toman decisiones. Desafortunadamente, con frecuencia se utilizan de manera incorrecta. Como resultado, los responsables de la toma de decisiones a menudo hacen predicciones inexactas y toman decisiones menos que deseables. Con la esperanza de que los evite, mencionaremos los errores más comunes cometidos en el uso de regresión y correlación.

Extrapolación más allá del rango de los datos observados

Límites específicos del rango para el que se cumple la ecuación de regresión

Un error común es suponer que la línea de estimación puede aplicarse en cualquier intervalo de valores. Los administradores de hospitales pueden utilizar adecuadamente el análisis de regresión para predecir la relación entre costos por cama y niveles de ocupación para varios niveles. Algunos administradores, sin embargo, utilizan incorrectamente la misma ecuación de regresión para predecir los costos por cama para niveles de ocupación que son significativamente más altos que los empleados para estimar la línea de regresión. Aun cuando una relación se cumpla para el intervalo de puntos de la muestra, puede existir una relación completamente distinta para un intervalo diferente. Como resultado, estas personas toman decisiones sobre un conjunto de costos y encuentran que cambian drásticamente al incrementarse la ocupación (debido a factores como los costos de tiempos extra y limitaciones de capacidad). Recuerde que **una ecuación de estimación es válida sólo para el mismo rango dentro del cual se tomó la muestra inicialmente.**

Causa y efecto

Los análisis de regresión y correlación no determinan la causa y el efecto

Otro error que podemos cometer al utilizar el análisis de regresión es suponer que un cambio en una variable es "ocasionado" por un cambio en la otra variable. Como se vio, **los análisis de regresión y correlación no pueden, de ninguna manera, determinar la causa y el efecto.** Si decimos que existe una correlación entre las calificaciones de los estudiantes en la universidad y sus ingresos anuales cinco años después de graduarse, no estamos diciendo que uno ocasiona al otro. Más bien, otros factores pueden ser la causa de ambos, como los antecedentes sociológicos, las actitudes paternas, la calidad de los profesores, la efectividad del proceso de entrevista para el trabajo y las condiciones económicas de los padres, por nombrar sólo unos cuantos factores potenciales.

Hemos utilizado extensamente el ejemplo relativo a los gastos de investigación y desarrollo y las ganancias anuales para ilustrar diversos aspectos del análisis de regresión. Pero, en realidad, es altamente improbable que las ganancias de un año dado estén ocasionadas por los gastos de ID en ese año. Ciertamente, sería temerario que el vicepresidente de ID sugiriera al director general que las ganancias podrían incrementarse de inmediato simplemente incrementando los gastos de ID. Particularmente en las industrias de alta tecnología, la actividad de ID puede usarse para explicar ganancias, pero una forma mejor de hacerlo sería predecir ganancias actuales en términos de gastos anteriores en investigación y desarrollo, así como en términos de condiciones económicas, dólares gastados en publicidad y otras variables. Esto puede hacerse utilizando las técnicas de regresión múltiple que se analizarán en el siguiente capítulo.

Uso de tendencias anteriores para estimar tendencias futuras

Las condiciones cambian e invalidan la ecuación de regresión

Debemos reevaluar los datos históricos que se usarán para estimar la ecuación de regresión. Las condiciones pueden cambiar y violar una o más de las suposiciones de las cuales depende nuestro análisis de regresión. Antes en este capítulo, hicimos notar que se supuso que la varianza de la perturbación o variación e alrededor de la media es constante. En muchas situaciones, sin embargo, esta varianza cambia de un año a otro.

Los valores de las variables cambian con el tiempo

Otro error que puede surgir del uso de datos históricos se refiere a la dependencia de algunas variables en el tiempo. Supongamos que una compañía utiliza el análisis de regresión para determinar la relación entre el número de empleados y el volumen de producción. Si las observaciones usadas en el análisis se remontan a varios años, la recta de regresión resultante puede estar demasiado inclinada porque puede no reconocer el efecto de los cambios en la tecnología.

Interpretación errónea de los coeficientes de correlación y determinación

Mala interpretación de r y r^2

Si $r = 0.6$, es incorrecto afirmar que la ecuación de regresión "explica" el 60% de la variación total en Y. Más bien, si $r = 0.6$, entonces r^2 debe ser $0.6 \times 0.6 = 0.36$. Sólo el 36% de la variación total se explica por la recta de regresión.

El coeficiente de determinación se malinterpreta si usamos r^2 para describir el porcentaje de cambio en la variable dependiente *ocasionado* por un cambio en la variable independiente. Esto es incorrecto porque r^2 es una medida sólo de qué tan bien una variable describe a la otra, *no* de qué tanto cambio en una variable es originado por la otra variable.

Descubrimiento de relaciones cuando no existen

Relaciones que no tienen un vínculo común

Al aplicar el análisis de regresión, la gente algunas veces encuentra una relación entre dos variables que, de hecho, no tienen un vínculo común. Aun cuando una variable no "ocasiona" un cambio en la otra, piensan que debe haber algún factor común a ambas variables. Sería posible, por ejemplo, encontrar una relación estadística entre una muestra aleatoria del número de millas por galón consumidas por ocho carros distintos y la distancia de la tierra a cada uno de los otros ocho planetas. Pero dado que no existe en absoluto un vínculo común entre la distancia recorrida por galón y la distancia a otros planetas, esta "relación" no tendría sentido.

Descubrimiento de cosas que no existen

A este respecto, si uno tuviera que desarrollar un gran número de regresiones entre muchos pares de variables, probablemente sería posible obtener algunas "relaciones" sugeridas bastante interesantes. Tal vez fuera posible, por ejemplo, encontrar una relación estadística entre su ingreso y la cantidad de cerveza consumida en Estados Unidos, o incluso entre la longitud de un tren (en carros) y el clima. Pero en ninguno de estos casos existe un factor común a ambas variables; por tanto, tales "relaciones" carecen de sentido. Como en la mayor parte de otras situaciones estadísticas, se requiere el conocimiento de las limitaciones inherentes a la técnica que se está empleando *además* de una gran dosis de sentido común para evitar llegar a conclusiones injustificadas.

SUGERENCIAS Y SUPOSICIONES

Advertencia: los administradores inteligentes *deben* poder razonar para llegar a una conexión de sentido común entre dos variables aun antes de realizar el análisis de regresión sobre esas variables. Pero las regresiones de computadora para bases de datos grandes, en ocasiones dan resultados sorprendentes en términos de relaciones no esperadas. Eso no invalida para nada el sentido común; lo que sugiere es que esos mismos administradores inteligentes prueben de nuevo estas "sorpresas" con una nueva muestra para ver si la relación "sorprendente" continúa siendo cierta. Sugerencia: piense que lo que *podría* tener entre manos es un problema de datos, no uno que contradice el sentido común.

Ejercicios 12.5

- **12-41** Explique por qué una ecuación de estimación es válida sólo en el intervalo de valores usados para su desarrollo.
- **12-42** Explique la diferencia entre el coeficiente de determinación y el coeficiente de correlación.
- **12-43** ¿Por qué debemos ser cautos al usar datos históricos para predecir tendencias futuras?
- **12-44** ¿Por qué no debemos atribuir causalidad en una relación aun cuando exista una fuerte correlación entre las variables o eventos?

Estadística en el trabajo

Loveland Computers

Caso 12: Regresión y correlación simples Loveland Computers estaba operando su línea de producción más seguido para ensamblar computadoras a partir de componentes ya disponibles, debido al crecimiento de la demanda de computadoras de alto rendimiento. Walter Azko tenía muy claro que esto era sólo ensamble, no "fabricación real". A menudo bromeaba que la única parte exclusiva de Loveland Computers era la base plástica para el teclado, adornada con el logotipo de Loveland (la silueta de las Rocallosas, justo como se ve desde la ventana de la oficina de Walt). La base consta de dos partes que embonan a presión. Y ése era el siguiente problema canalizado a Lee Azko. Nancy Rainwater, la supervisora de producción, explicaba sus frustraciones a Lee.

"Cuando empezamos a ensamblar este modelo el verano pasado, las bases del teclado parecían embonar perfectamente. Ahora tenemos que rechazar muchas de ellas porque las pequeñas pestañas que sostienen la parte alta de la base se rompen cuando el operador las presiona para unirlas. Cuando eso sucede, tenemos que tirar ambas piezas. No contamos con forma de reciclar ese tipo de plástico, y no parece correcto estar mandando todo eso al relleno sanitario, por no mencionar lo que le está haciendo a nuestros costos.

"Hablé con compras e hice que Tyronza Wilson inspeccionara las bases al recibirlas. Las medidas de las pestañas cumplen exactamente con las especificaciones, y la compañía de plásticos que nos las fabrica hizo cierto trabajo de laboratorio. Dicen que no encontraron defectos en el plástico que están usando.

"Noté que teníamos más roturas temprano en la mañana, así que me pregunté si esto sucedía simplemente porque la gente no tenía cuidado en la línea. Incluso llegué a preguntarme si no sería porque los empleados no tuvieran la capacitación adecuada; pero el hecho es que esta gente tiene más experiencia ahora que el verano pasado, realmente no hemos tenido mucha rotación de personal.

"Tyronza se preguntaba si esto sucede porque el plástico está demasiado frío. Eso lo explicaría todo si hubiera más defectos en invierno. Pero el almacén tiene un par de calentadores, así que no estoy segura de que eso sea correcto. Y yo realmente no puedo andar con un termómetro, verificando la temperatura de cada juego de partes para las bases antes de enviarlas a la línea, ¿o sí?".

"Tal vez haya otra forma de resolver esto", dijo Lee, recordando que había sido bastante simple obtener estadísticas climáticas del Servicio Meteorológico Nacional. "Registraste el número de bases desechadas por cada día de operación de la línea de producción, ¿o no?"

Preguntas de estudio: ¿Cómo investigaría Lee la relación entre el clima y el problema con las bases de plástico? ¿"Probará" esto que la explicación de Tyronza es correcta?

Ejercicio de base de datos computacional

HH Industries

Hal buscó a Laurel poco después de su regreso de las Rocallosas. "Realmente te ves descansada", comentó. "Probablemente a mí también me vendrían bien unas vacaciones, pero me temo que tendré que esperar un tiempo. ¡La época más atareada del año está por llegar! A propósito, quisiera que vieras algo por mí. Estamos en posición de contratar personal adicional para el almacén, tanto aquí como en nuestras sucursales, sobre todo para tareas 'no calificadas' como envío, recepción, empaque, despacho de pedidos, etc. Lo que quisiera saber es si hay alguna 'fórmula' que la estadística nos pudiera demostrar que es mejor que otras. Hemos tenido resultados mezclados en el pasado. Resulta caro en estos días contratar y capacitar gente, y nuestros costos de personal se reducen considerablemente cuando reducimos la rotación. ¿Crees poder ayudarnos?"

"Suena como que se pudiera aplicar un poco de análisis de regresión", dijo Laurel. "Hablaré con Gary, ya que estamos hablando de su personal, y veré qué puedo obtener."

Hal sonrió. “Magnífico. Mi secretaria, Mary, tiene todos los archivos de personal sobre empleados actuales y anteriores. Sé que no somos una compañía enorme, pero al menos tenemos *algunos* datos puntuales para que analices.”

Laurel se dirigió al almacén para ver a Gary. “Te haré saber qué encuentro”, le dijo por encima del hombro.

Gary, ocupado con un embarque que acababa de llegar, no tenía mucho tiempo para platicar. Después de fijar una cita para la siguiente tarde, logró darle una idea de dónde empezar a Laurel. “Hemos tenido éxito con nuestro programa de contratación de jubilados. Son trabajadores estables, contentos de poder estar ocupados en algo, ¡y hay muchos aquí en Florida! Tal vez la edad podría ser el tipo de característica que buscas. Sin embargo, te prometo que antes de mañana pensaré más en ello.”

“Gracias”, dijo Laurel. “Y siento haberte interrumpido.”

“No hay problema”, Gary le sonrió brevemente y regresó a su tarea.

Después de recabar los datos adecuados con Mary, Laurel se dirigió a su computadora. Para evaluar con precisión el factor de “periodo de empleo”, sabía que lo más probable era que tuviera que usar los datos de exempleados. Sin embargo, unos cuantos empleados actuales del almacén tenían varios años con la compañía, y sentía que eran bastante importantes como para incluirlos. Después de un corto estudio de la información disponible, decidió incluir a los empleados actuales con cinco años o más de servicio.

1. Realice una regresión lineal de mínimos cuadrados sobre los datos proporcionados en los archivos CH12.XXX del CD que acompaña al libro. ¿Cuál es el error estándar de la estimación? Suponiendo distribuciones normales alrededor de cada valor estimado y varianzas iguales en cada punto, calcule un intervalo de predicción aproximado del 95.5% (± 2 errores estándar) para la duración de empleo de un empleado potencial de 25 años de edad. Haga el mismo cálculo para un empleado potencial de 65 años. Dada esta información solamente, ¿podemos hacer algunas recomendaciones respecto a qué persona contratar?
2. ¿Cuáles son los coeficientes de determinación y correlación para la duración de empleo (en meses) contra edad al contratarlo (en años)?
3. Gary siempre ha sentido que (siendo los otros factores iguales) cada año adicional de edad de un empleado potencial corresponde a un mes más de empleo en HH Industries. Pruebe la hipótesis de que la pendiente de la línea de regresión de población es 1.0 al nivel de significancia del 10%.

Del libro de texto al mundo real

Aplicaciones de métodos estadísticos al fútbol americano

Aunque el uso de los métodos estadísticos es más común en las áreas de negocios, también tienen una importante función en el mundo de los deportes. Para los no iniciados, el fútbol americano se caracteriza por contrincantes fuertemente acorazados atacándose a toda velocidad y tirándose mutuamente al suelo. Debajo de esta apariencia de Neanderthal, radica un juego de notable complejidad, donde la estadística desempeña un papel importante. Los entrenadores usan las estadísticas para idear estrategias para juegos específicos, y los periodistas de deportes para clasificar equipos y predecir resultados de partidos.

Implicaciones estratégicas Como cada juego de fútbol americano empieza con un reinicio, los jugadores ofensivos y defensivos tienen una oportunidad de alinearse contra sus oponentes; por tanto, la planeación estratégica es esencial. Las estadísticas típicas incluyen la distancia promedio ganada por carrera, el porcentaje de pases completos, la distancia promedio ganada por pases completos recibidos con éxito, la distancia promedio al patear el balón, el número de veces que el balón se deja caer y el número de pases interceptados. Estas estadísticas se llevan por individuo y por equipo. En los años sesenta, los Vaqueros de Dallas, de la Liga Nacional (NFL), comenzaron a utilizar datos de juegos individuales para identificar las tendencias mostradas por los equipos ofensivos contrarios y para eliminar tendencias visibles en sus propios jugadores ofensivos. Como uno podría sospechar, los Vaqueros fueron uno de los equipos más exitosos durante ese periodo. Al revisar las estadísticas de sus contrincantes, el cuerpo técnico espera encontrar tendencias donde el equipo oponente use de manera consistente una jugada o una formación en particular. Una vez identificado, los jugadores defensivos pueden alinearse para detener la jugada esperada. Hoy en día, los 28 equipos de la NFL utilizan métodos estadísticos para determinar jugadas defensivas y establecer estrategias ofensivas. Las estadísticas individuales también desempeñan un importante papel en el procedimiento de contratación de jugadores.

Clasificación y predicciones Las predicciones automatizadas se han asociado con el fútbol americano durante más de 50 años. El “sistema” Williamsen fue ampliamente publicado en periódicos durante los años treinta. Williamsen utilizaba una técnica de mínimos cuadrados para clasificar equipos universitarios y predecir resultados. Las encuestas de servicio cablegráfico de agencias periodísticas nacionales comenzaron en 1936 después de la popularidad de los datos de Wi-

lliamsen. Estas encuestas, que clasifican a los 20 equipos universitarios más importantes, se siguen utilizando actualmente. Raymond Stefani, profesor de ingeniería eléctrica en la Universidad del Estado de California, proporcionó predicciones semanales sobre más de 11,000 juegos, comenzando con la temporada 1970-1971 y finalizando con la de 1980-1981; para ello utilizó un procedimiento de mínimos cuadrados. El empleo de mínimos cuadrados permitió a Stefani predecir el equipo ganador correcto en el 70% de esos juegos.

Fuente: Raymond T. Stefani, "Applications of Statistical Methods to American Footba1l", en *Journal of Applied Statistics* 14(1) (1987): 61-73.

Repaso del capítulo

● Términos introducidos en el capítulo 12

Análisis de correlación Técnica para determinar el grado en el que las variables se relacionan linealmente.

Coeficiente de correlación Raíz cuadrada del coeficiente de determinación. Su signo indica la dirección de la relación entre dos variables, directa o inversa.

Coeficiente de determinación Medida de la proporción de variación en *Y*, la variable dependiente, que explica la recta de regresión, esto es, la relación de *Y* con la variable independiente.

Diagrama de dispersión Gráfica de puntos en una cuadrícula; las coordenadas *X* y *Y* de cada punto corresponden a las dos mediciones hechas sobre un elemento particular de la muestra; el patrón de puntos ilustra la relación entre las dos variables.

Ecuación de estimación Fórmula matemática que relaciona la variable desconocida con las variables conocidas en el análisis de regresión.

Error estándar de la estimación Medida de la confiabilidad de la ecuación de estimación, que indica la variabilidad de los puntos observados alrededor de la recta de regresión, esto es, de qué manera los valores observados difieren de sus valores pronosticados sobre la recta de regresión.

Error estándar del coeficiente de regresión Medida de la variabilidad del coeficiente de regresión de la muestra alrededor del coeficiente de regresión verdadero de la población.

Método de mínimos cuadrados Técnica para ajustar una línea recta a través de un conjunto de puntos de tal manera que la suma de los cuadrados de las distancias verticales de los *n* puntos a la recta se minimiza.

Ordenada *Y* Constante para cualquier línea recta dada cuyo valor representa el valor de la variable *Y* cuando el valor de la variable *X* es 0.

Pendiente Constante para cualquier línea recta dada cuyo valor representa cuánto cambia la variable dependiente con un cambio de una unidad de la variable independiente.

Recta de regresión Una línea ajustada a un conjunto de datos para estimar la relación entre dos variables.

Regresión Proceso general para predecir una variable a partir de otra mediante medios estadísticos utilizando datos históricos.

Regresión múltiple Proceso estadístico mediante el cual se utilizan varias variables para predecir otra variable.

Relación curvilínea Asociación entre dos variables que se describe por una línea curva.

Relación directa Relación entre dos variables en donde, al aumentar el valor de la variable independiente, aumenta el valor de la variable dependiente.

Relación inversa Relación entre dos variables en donde, al aumentar la variable independiente, la variable dependiente disminuye.

Relación lineal Tipo particular de asociación entre dos variables que puede describirse matemáticamente mediante una línea recta.

Variable dependiente La variable que tratamos de predecir en el análisis de regresión.

Variables independientes Variable o variables conocidas en el análisis de regresión.

● Ecuaciones introducidas en el capítulo 12

■ **12-1**

$$Y = a + bX$$

Ecuación de una *línea recta*, donde la variable dependiente *Y* está "determinada" por la variable independiente *X*. La *a* se llama *ordenada Y* porque su valor es el punto en el cual la recta cruza el eje *Y* (el eje vertical). La *b* es la *pendiente* de la recta, esto es, dice cuánto cambia la variable dependiente *Y* con cada

cambio unitario de la variable independiente X. Tanto a como b son constantes numéricas, ya que para una línea recta dada, sus valores no cambian.

■ **12-2**
$$b = \frac{Y_2 - Y_1}{X_2 - X_1}$$

Para calcular la constante numérica b para una recta dada, encuentre el valor de las coordenadas, X y Y, para dos puntos que están en la recta. Las coordenadas para el primer punto son $(X_1, Y_1,)$ y el segundo punto (X_2, Y_2). Recuerde que b es la pendiente de la recta.

■ **12-3**
$$\hat{Y} = a + bX$$

En el análisis de regresión, $\hat{Y}$ (Y *gorro*) simboliza los valores individuales de Y de los puntos *estimados*, esto es, los puntos que están en la línea de estimación. En consecuencia, la ecuación 12-3 es la ecuación para la línea de estimación.

■ **12-4**
$$b = \frac{\Sigma XY - n\overline{X}\,\overline{Y}}{\Sigma X^2 - n\overline{X}^2}$$

La ecuación nos permite calcular la *pendiente de la recta de regresión de mejor ajuste* para cualquier conjunto de puntos de dos variables. Introdujimos dos nuevos símbolos en esta ecuación, $\overline{X}$ y $\overline{Y}$, que representan las medias de los valores de la variable independiente y la variable dependiente, respectivamente. Además esta ecuación contiene a n que, en este caso, es el número de puntos para los cuales se ajusta la recta de regresión.

■ **12-5**
$$a = \overline{Y} - b\overline{X}$$

Con esta fórmula podemos calcular la *ordenada Y de la recta de regresión de mejor ajuste* para un conjunto de puntos de dos variables.

■ **12-6**
$$s_e = \sqrt{\frac{\Sigma(Y - \hat{Y})^2}{n - 2}}$$

El *error estándar de la estimación*, s_e, mide la variabilidad o dispersión de los valores observados alrededor de la recta de regresión. En efecto, indica la confiabilidad de la ecuación de estimación. El denominador es $n - 2$ porque perdemos 2 grados de libertad (para los valores a y b) al estimar la recta de regresión.

■ **12-7**
$$s_e = \sqrt{\frac{\Sigma Y^2 - a\Sigma Y - b\Sigma XY}{n - 2}}$$

Como la ecuación 12-6 requiere cálculos tediosos, los estadísticos han ideado este *método corto para encontrar el error estándar de la estimación*. Al calcular los valores para b y a, ya se calcularon las cantidades de la ecuación 12-7, excepto ΣY^2, es muy sencillo obtener.

■ **12-8** Variación de los valores de Y alrededor de la recta de regresión $= \Sigma(Y - \hat{Y})^2$

La variación de los valores de Y en un conjunto de datos alrededor de la recta de regresión ajustada es una de dos cantidades a partir de las cuales se desarrolla el coeficiente de determinación de la muestra. La ecuación 12-8 indica cómo medir esta dispersión particular, que es la porción *no explicada* de la variación total de los valores de Y.

■ **12-9** Variación de los valores de Y alrededor de su propia media $= \Sigma(Y \text{ y } \overline{Y})^2$

Esta fórmula mide la *variación total* de un conjunto completo de valores de Y, esto es, la variación de estos valores de Y alrededor de su propia media.

■ **12-10**
$$\tilde{r}^2 = 1 - \frac{\Sigma(Y - \hat{Y})^2}{\Sigma(Y - \overline{Y})^2}$$

El *coeficiente de determinación de la muestra*, r^2, da la fracción de la variación total de Y que explica la recta de regresión. Es una importante medida del grado de asociación entre X y Y. Si el valor de r^2 es $+1$, entonces la recta de regresión es un estimador perfecto. Si $r^2 = 0$, no existe correlación entre X y Y.

■ **12-11**
$$r^2 = \frac{a\Sigma Y + b\Sigma XY - n\overline{Y}^2}{\Sigma Y^2 - n\overline{Y}^2}$$

Ésta es una ecuación de método corto para calcular r^2.

■ **12-12**
$$r = \sqrt{r^2}$$

El *coeficiente de correlación de la muestra* se denota por r y se encuentra tomando la raíz cuadrada del coeficiente de determinación de la muestra. Es una segunda medida (además de r^2) que podemos utilizar para describir qué tan bien una variable explica a otra. El signo de r es igual al signo de b; indica la dirección de la relación entre las dos variables X y Y.

■ **12-13**
$$Y = A + BX$$

Toda *recta de regresión de la población* tiene la forma de la ecuación 12-13, donde A es la intersección Y para la población, y B es la pendiente.

■ **12-13a**
$$Y = A + BX + e$$

Como no todos los puntos individuales de un población están en la recta de regresión de la población, los puntos *individuales* satisfacen la ecuación 12-13a, en donde e es una variación aleatoria respecto a la recta de regresión de la población. En promedio, e es igual a cero, porque las variaciones arriba de la recta de regresión se cancelan con las variaciones que se encuentran abajo de ella.

■ **12-14**
$$s_b = \frac{s_e}{\sqrt{\Sigma X^2 - n\bar{X}^2}}$$

Al manejar una muestra, podemos usar esta fórmula para obtener el *error estándar del coeficiente de regresión, b*.

■ **12-15**
$$t = \frac{b - B_{H_0}}{s_b}$$

Una vez calculado s_b con la ecuación 12-14, podemos usar esta ecuación para estandarizar el valor observado del coeficiente de regresión. Después realizamos la prueba de hipótesis comparando este valor estandarizado con el o los valores críticos de la tabla 2 del apéndice.

● Ejercicios de repaso

■ **12-45** Un consultor está interesado en el grado de precisión con que un nuevo índice de desempeño laboral mide lo que es importante para una corporación. Una forma de verificarlo es analizar la relación entre el índice de evaluación del trabajo y el salario de un empleado. Se tomó una muestra de ocho empleados y se recabó información del salario (en miles de dólares) y el índice de evaluación del trabajo (1 a 10, donde 10 es la mejor calificación).

Índice de evaluación del trabajo (*X*)	9	7	8	4	7	5	5	6
Salario (*Y*)	36	25	33	15	28	19	20	22

a) Desarrolle la ecuación de estimación que mejor describa estos datos.
b) Calcule el error estándar de la estimación, s_e, para estos datos.
c) Calcule el coeficiente de determinación de la muestra, r^2, para estos datos.

■ **12-46** La Stork Foundation desea mostrar con estadísticas que, contrariamente a la creencia popular, las cigüeñas sí traen a los bebés. Para esto ha recolectado datos sobre el número de cigüeñas y el número de bebés (ambos en miles) en varias ciudades grandes de Europa central.

Cigüeñas	27	38	13	24	6	19	15
Bebés	35	46	19	32	15	31	20

a) Calcule el coeficiente de determinación de la muestra y el coeficiente de correlación de la muestra para estos datos.
b) ¿Contradijo la ciencia estadística la creencia popular?

■ **12-47** (Llene los espacios en blanco.) Los análisis de regresión y correlación tratan la _____ entre variables. El análisis de regresión, mediante ecuaciones _____, nos permite _____ una variable desconocida a partir de un conjunto de variables conocidas. La variable desconocida se llama variable _____; las variables conocidas se denominan variables _____. La correlación entre dos variables indica el _____ de la relación lineal entre ellas y por tanto da una idea de qué tan bien el _____ de regresión describe la relación entre las variables.

■ **12-48** Calcule el coeficiente de determinación de la muestra y el coeficiente de correlación de la muestra para el ejercicio 12-14.

Business Week y *U.S. News & World Report* publican clasificaciones de las mejores 20 escuelas de administración. La clasificación global del *Business Week* se basa en clasificaciones obtenidas de estudiantes y compañías que reclutan maestros en administración. Junto con las clasificaciones, las publicaciones reportan información sobre el costo de obtener una maestría y los salarios iniciales promedio de los graduados. Utilice los datos de la tabla MR12-1 para responder los ejercicios 12-49 a 12-52.

Tabla ER12-1

Estudios de clasificación de escuelas de administración

Escuela	Clasificación de 1992 *BW*	Clasificación de 1992 *USN&WR*	Clasificación de *BW* por estudiantes	Clasificación de *BW* por compañías	Costo	Salario inicial
Northwestern	1	4	3	1	37,600	70,200
Chicago	2	6	10	4	38,500	68,600
Harvard	3	2	12	3	37,100	84,960
Wharton	4	3	15	2	37,600	72,200
Michigan	5	7	9	6	37,200	58,110
Dartmouth	6	10	1	12	37,500	74,260
Stanford	7	1	5	7	38,480	82,860
Indiana	8	18	6	8	24,600	49,070
Columbia	9	8	18	5	38,000	66,620
North Carolina	10	16	8	11	17,360	55,500
Virginia	11	11	2	15	28,500	65,280
Duke	12	9	7	14	37,000	59,870
MIT	13	5	14	10	39,000	73,000
Cornell	14	12	4	17	37,000	59,940
NYU	15	17	16	13	36,100	56,730
UCLA	16	14	11	16	22,500	64,540
Carnegie-Mellon	17	15	23	9	37,200	56,980
Berkeley	18	13	13	19	15,400	65,500
Vanderbilt	19	19	19	20	35,000	47,320
Washington	20	20	24	18	33,500	48,200

Fuente: Adaptado de Business Week *(26 de octubre de 1992): 60 y* U.S. News & World Report *(23 de marzo de 1992): 66.*

12-49 Trace un diagrama de dispersión de la clasificación *USN&WR* contra el costo del grado de la maestría. ¿Parece que las escuelas más caras obtienen mejores clasificaciones? Calcule el coeficiente de correlación de muestra entre estas dos variables.

12-50 ¿Existe una retribución por gastar más en una maestría? Grafique un diagrama de dispersión del salario inicial contra el costo. Ajuste una ecuación de regresión a los datos y pruebe las hipótesis apropiadas respecto a su pendiente.

12-51 ¿Los graduados de escuelas con clasificación más alta obtienen salarios iniciales más altos? Trace un diagrama de dispersión de salario inicial contra la clasificación global de *Business Week*. Ajuste una ecuación de regresión a los datos y pruebe las hipótesis apropiadas respecto a su pendiente.

12-52 ¿Qué tan fuerte es la relación entre los salarios iniciales y las clasificaciones? Calcule los coeficientes de determinación de la muestra entre los salarios iniciales y las tres clasificaciones de *Business Week* (globales, por estudiantes y por compañías). ¿Cuáles de estas clasificaciones explican la mayor parte de la variación en salarios iniciales?

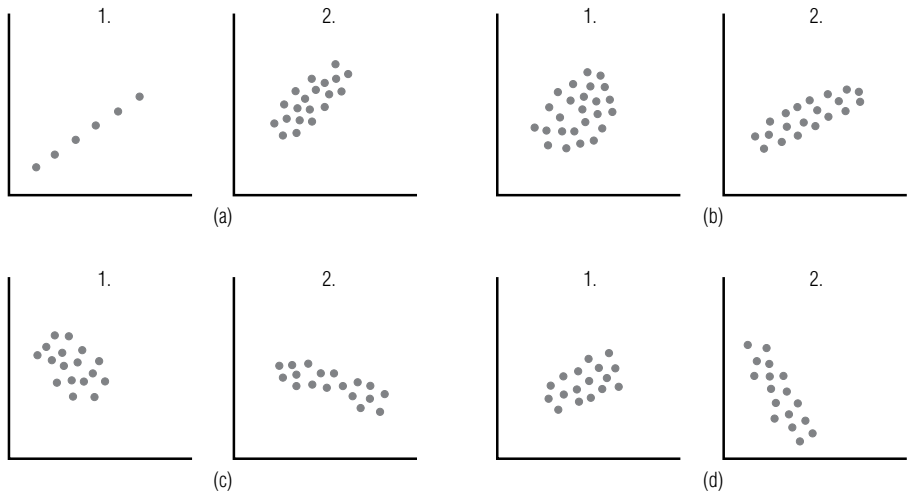
■ **12-53** "Nada triunfa como el éxito" es un antiguo adagio en el negocio de la publicidad. El presidente de una distribuidora de varias líneas de automóviles ha observado que los agentes de ventas que gana los bonos más altos al final de año son los que tienen mayor probabilidad de exceder su cuota de ventas el año siguiente (y ganar otro bono alto).

Bono el año pasado (miles de dólares)	7.8	6.9	6.7	6.0	6.9	5.2
Ventas arriba de cuota este año	64	73	42	49	71	46
Bono el año pasado (miles de dólares)	6.3	8.4	7.2	10.1	10.8	7.7
Ventas arriba de cuota este año	32	88	53	84	85	93

a) Desarrolle la recta que mejor se ajuste para describir estos datos.
b) Calcule el error estándar de la estimación para la relación.

c) Desarrolle un intervalo de confianza de aproximadamente el 90% para predecir las ventas arriba de la cuota para un miembro del personal que ganó un bono de $9,600 el año pasado.

■ **12-54** Para cada uno de los siguientes pares de diagramas diga cuál tiene un valor más alto de *r*, el coeficiente de correlación y cuál es el signo de *r*.



■ **12-55** Un gerente de operaciones está interesado en predecir los costos *C* (en miles de dólares) con base en la cantidad de materia prima de entrada *R* (en miles de libras) para un fabricante de pantalones de mezclilla. Si la pendiente es significativamente mayor que 0.5 en los siguientes datos muestrales, entonces algo marcha mal con el proceso de producción y la maquinaria de la línea de ensamble debe ajustarse. Al nivel de significancia de 0.05, ¿debe ajustarse la maquinaria? Establezca explícitamente las hipótesis y una conclusión.

C	10	7	5	6	7	6
R	25	20	16	17	19	18

■ **12-56** Calcule el coeficiente de determinación de la muestra y el coeficiente de correlación de la muestra para el ejercicio 12-13.

■ **12-57** No debemos extrapolar para predecir valores fuera del intervalo de datos usados al construir la recta de regresión. La razón (elija una):

a) La relación entre las variables puede no ser la misma para otros valores de las variables.
b) La variable independiente puede no tener el efecto causal sobre la variable dependiente para estos valores.
c) Los valores de las variables pueden cambiar con el tiempo.
d) Tal vez no exista un vínculo común para explicar la relación.

Utilice los datos de 50 áreas metropolitanas de Estados Unidos dados en la tabla MR11-2, al final del capítulo 11, para responder los ejercicios 12-58 a 12-60.

12-58 A menudo, quienes planean la comercialización deben estimar la demanda geográfica de un producto de una compañía. La demanda no depende sólo del número de personas de una comunidad, sino también de la cantidad de dinero que tienen para gastar. La revista *Sales & Marketing Management* utiliza los datos de censos de Estados Unidos para estimar el "ingreso de compra efectivo (ICE)" de hogares típicos estadounidenses en áreas metropolitanas del país. El ICE es la suma de sueldos y otros ingresos, menos impuestos y multas; en resumen, es una buena medida de lo que los economistas llaman "ingreso disponible". La cantidad total de dinero disponible para gastar en una comunidad es aproximadamente proporcional al producto del ICE por la población.

Calcule una nueva variable TD = (POB × ICE)/1,000. Calcule los coeficientes de determinación de la muestra entre VENTAS y POB y entre VENTAS y TD. ¿Cuál de estas variables explica una parte mayor de la variación en VENTAS?

12-59 Ajuste una ecuación de regresión que use SOLA para predecir el valor de VENTAS. Encuentre un intervalo de predicción del 90% para las ventas totales al menudeo en un área metropolitana que tiene 20,000 casas con una sola persona. ¿Hasta qué punto sería útil este resultado para una compañía de productos de consumo que está desarrollando una nueva línea de cenas congeladas individuales?

12-60 Suponga que desea saber si los negocios son mejores en comunidades con más gente mayor. Utilice una edad promedio para representar el número de personas mayores en cada área metropolitana, ajuste una ecuación de regresión para explicar VENTAS en términos de EDAD. ¿La pendiente de su regresión es significativamente mayor que cero? Con base en este análisis, ¿debe concluir que "los negocios *no son* mejores en comunidades con más personas mayores"? Explique su respuesta.

■ **12-61** Los economistas con frecuencia están interesados en estimar funciones de consumo, que se obtienen mediante la regresión del consumo Y sobre el ingreso X (para esta regresión, los economistas llaman a la pendiente la *propensión marginal al consumo*). Para una muestra de 25 familias, se calculó una pendiente de 0.87 y un error estándar del coeficiente de la pendiente de regresión de 0.035. Para esta muestra, ¿la propensión marginal a consumir disminuyó a menos que el estándar de 0.94? Utilice $\alpha = 0.05$. Establezca las hipótesis explícitas y una conclusión.

■ **12-62** A diferencia del coeficiente de determinación, el coeficiente de correlación (escoja la respuesta correcta):

a) Indica si la pendiente de la recta de regresión es positiva o negativa.
b) Mide la fuerza de asociación entre las dos variables de manera más exacta.
c) Nunca puede tener un valor absoluto mayor que 1.
d) Mide el porcentaje de varianza explicado por la recta de regresión.

■ **12-63** ¿Son importantes las calificaciones en la universidad para ganar un buen sueldo? Un estudiante de estadística para la administración tomó una muestra aleatoria de sueldos iniciales y promedios de calificaciones en la universidad de algunos de sus amigos recién graduados. Los datos son los siguientes:

Sueldo inicial (miles de dólares)	36	30	30	24	27	33	21	27
Promedio de calificaciones	4.0	3.0	3.5	2.0	3.0	3.5	2.5	2.5

a) Grafique estos datos.
b) Desarrolle la ecuación de estimación que mejor describa los datos.
c) Grafique la ecuación de estimación en el diagrama de dispersión del inciso a).

■ **12-64** Un arrendador está interesado en ver si las rentas de sus departamentos son las comunes. Para esto tomó una muestra aleatoria de 11 rentas y tamaños de departamentos en complejos de departamentos similares. Los datos son los siguientes:

Renta	230	190	450	310	218	185	340	245	125	350	280
Número de recámaras	2	1	3	2	2	2	2	1	1	2	2

a) Desarrolle la ecuación de estimación que mejor describa estos datos.
b) Calcule el coeficiente de determinación.
c) Pronostique la renta para un departamento de dos recámaras.

■ **12-65** Muchas compañías pequeñas compran publicidad sin analizar sus efectos. La "guerra de las hamburguesas" (rivalidad sustancial de precios entre compañías de comida rápida) ha reducido las ganancias de Ethiopian Burguers en Santa Cruz, California, una cadena regional pequeña. El gerente de mercadotecnia intenta demostrar que "hay que gastar dinero para ganar dinero". Gastar en publicidad en espectaculares, en su opinión, tiene resultados directos en las ventas. Se tienen registros de 7 meses:

Gasto mensual en espectaculares (miles de dólares)	25	16	42	34	10	21	19
Rendimiento de las ventas mensuales (miles de dólares)	34	14	48	32	26	29	20

a) Desarrolle la ecuación de estimación que mejor describa estos datos.
b) Calcule el error estándar de la estimación para esta relación.
c) Para un mes con gastos de $28,000 en espectaculares, desarrolle un intervalo de confianza del 95% para las ventas mensuales esperadas ese mes.

12-66 En 1992, las ventas totales de cereales para desayuno en Estados Unidos se estimaron en $3.842 miles de millones. Considere la siguiente información de los 10 cereales más vendidos. Encuentre la ecuación de mínimos cuadrados que usa el precio promedio al menudeo para predecir las participaciones en el mercado. ¿Cuál de las siguientes tres generalizaciones describe mejor la relación entre estas dos variables?

a) Un precio menor incrementa las ventas.
b) Un porcentaje de mercado mayor significa que se puede cobrar un precio más alto.
c) El porcentaje de mercado no parece depender del precio.

Compañía	Cereal	Porcentaje de mercado	Volumen (millones de dólares)	Precio promedio al menudeo
General Mills	Cheerios	4.58	175.96	$2.18
Kellogg's	Frosted Flakes	4.08	156.75	$2.83
General Mills	Honey Nut Cheerios	3.28	126.02	$2.99
Kellogg's	Rice Krispies	2.99	114.88	$1.94
Kellogg's	Corn Flakes	2.97	114.11	$1.47
Kellogg's	Raisin Bran	2.77	106.42	$2.74
Kellogg's	Frosted Mini-Wheats	2.77	106.42	$2.91
Kellogg's	Froot Loops	2.33	89.52	$2.64
General Mills	Lucky Charms	1.85	71.08	$3.15
General Mills	Total	1.84	70.69	$2.86

Fuente: Richard Gibson, "There Is No Way to Sugarcoat This News: Prices of Breakfast Cereals Are Going Up", The Wall Street Journal *(21 de enero de 1993): B1.*

■ **12-67** La autoridad aeronáutica estadounidense realizó un estudio de operaciones de aerolíneas, en 18 compañías, que reveló que la relación entre el número de pilotos empleados y el número de aviones en servicio tenía una pendiente de 4.3. Estudios anteriores indicaban que la pendiente de esta relación era 4.0. Si se calculó que el error estándar del coeficiente de pendiente de regresión es 0.17, ¿hay razones para creer, a un nivel de significancia de 0.05, que la pendiente verdadera ha cambiado?

■ **12-68** Dave Proffitt, estudiante de segundo año de la maestría en administración, elabora un estudio de compañías que entran a la bolsa de valores por primera vez. Tiene curiosidad por ver si existe o no una relación significativa entre el tamaño de la oferta (en millones de dólares) y el precio por acción.

a) Dados los siguientes datos, desarrolle la ecuación de estimación que mejor ajuste los datos.

Tamaño (millones de dólares)	Precio (dólares)
108.00	12.00
4.40	4.00
3.50	5.00
3.60	6.00
39.00	13.00
68.40	19.00
7.50	8.50
5.50	5.00
375.00	15.00
12.00	6.00
51.00	12.00
66.00	12.00
10.40	6.50
4.00	3.00

b) Calcule el coeficiente de determinación de la muestra. ¿Debe Dave usar esta ecuación de regresión para pronosticar o debe buscar en otra parte variables explicativas adicionales?

■ **12-69** Un fabricante de teléfonos celulares está probando dos tipos de baterías para ver cuánto duran con una utilización normal. La siguiente tabla contiene los datos provisionales:

	Vida aproximada (meses)	
Horas de uso diario	Litio	Alcalina
2.0	3.1	1.3
1.5	4.2	1.6
1.0	5.1	1.8
0.5	6.3	2.2

a) Desarrolle dos ecuaciones de estimación lineales, una para pronosticar la vida del producto basada en el uso diario con las baterías de litio y otra para las baterías alcalinas.

b) Encuentre un intervalo de confianza para la estimación del 90% para la vida (en meses) con 1.25 horas de uso diario, para cada tipo de batería. ¿Puede la compañía asegurar algo respecto a qué batería proporciona la vida más larga según estos números?

■ **12-70** Se ha propuesto un estudio para investigar la relación entre el peso al nacer de bebés varones y su estatura de adultos. Usando los siguientes datos, desarrolle la ecuación de estimación de mínimos cuadrados. ¿Qué porcentaje de la variación en la altura de adultos explica esta recta de regresión?

Peso al nacer	Estatura de adulto
5 lb, 8 oz	5′9″
7 lb	6′
6 lb, 4 oz	5′6″
7 lb, 8 oz	5′11″
8 lb, 2 oz	6′1″
6 lb, 12 oz	5′10″

■ **12-71** Muchos estudiantes universitarios se cambian de universidad el verano anterior al tercer año. Para ayudar a evaluar el potencial académico de quienes se cambian, Barbara Hoopes, la directora de admisiones del Piedmont College, realiza un análisis que compara los promedios globales de los estudiantes (PG) durante sus primeros dos años de universidad con los PG de sus últimos dos años, después del cambio. Usando los siguientes datos:

PG de primero y segundo año	1.7	3.5	2.3	2.6	3.0	2.8	2.4	1.9	2.0	3.1
PG de penúltimo y último año	2.4	3.7	2.0	2.5	3.2	3.0	2.5	1.8	2.7	3.7

a) Calcule la ecuación de estimación de mínimos cuadrados que debe usar Hoopes para predecir el PG el tercero y último año de licenciatura de los estudiantes que se cambian al Piedmont College
b) Hoopes no admitirá solicitantes de cambio de penúltimo año a menos que los intervalos de predicción del 90% para sus PG de penúltimo/último año definitivamente arriba de 2.0. ¿Admitirá un solicitante de cambio con un PG de primero/segundo de 2.5?

■ **12-72** Los salarios de muchos funcionarios públicos son menores que los que podrían tener con trabajos similares en la industria privada. *The Wall Street Journal* publicó los salarios de 10 procuradores generales y los comparó con el salario típico de un abogado al entrar a trabajar, en el mismo estado. Al responder a las siguientes preguntas, suponga que los salarios al entrar a trabajar son un buen indicador de la tasa en el mercado para los abogados.

	Procurador general	Abogado principiante
Vermont	61,025	26,520
Wyoming	75,000	31,500
Massachusetts	80,000	25,000
Pennsylvania	84,000	33,819
Georgia	90,000	35,880
Washington	92,000	30,000
California	102,000	38,400
Illinois	105,387	27,048
Nueva York	110,000	33,922
Michigan	111,200	35,182

Fuente: "Paying States' Attorneys General", The Wall Street Journal *(24 de julio de 1995): B8.*

a) ¿Varía el salario ofrecido al procurador general de acuerdo con la tasa para los abogados en cada estado? Pruebe, para $\alpha = 0.05$, si la pendiente de la regresión ajustada es significativamente diferente de 0.
b) ¿Qué proporción de variación en los salarios del procurador general se explica por la tasa para los abogados en el mercado lucrativo?
c) Si un procurador general desea elevar el ingreso en todo el estado para los abogados, ¿ayudaría presionar por un aumento en el salario del procurador general? ¿Por qué sí o por qué no?

■ **12-73** Los costos de los viajes de negocios varían mucho entre las ciudades más importantes de Estados Unidos, como se muestra en la siguiente tabla. Un interventor corporativo intenta establecer tasas de *viáticos* que tomen en cuenta esta variación. ¿Debe el interventor considerar los costos tanto de renta de autos como de hoteles, o los costos de hoteles proporcionan suficiente información para calcular las tasas? (*Sugeren-*

cia: ajuste una regresión usando los costos de la renta de autos para explicar los costos de los hoteles. Después observe r^2.)

	Hotel (dólares)	Renta de auto/día (dólares)
Atlanta	121	54
Boston	199	50
Chicago	159	62
Cleveland	129	52
Dallas	117	44
Denver	92	35
Detroit	102	60
Houston	92	70
Los Angeles	122	51
Miami	111	32
Minneapolis	107	57
Nueva Orleans	116	42
Nueva York	197	60
Orlando	95	36
Phoenix	85	37
Pittsburgh	122	46
St. Louis	115	66
San Francisco	155	52
Seattle	125	45
Washington, D.C.	145	53

Fuente: "Dow Jones Travel Index", The Wall Street Journal *(4 de agosto de 1995): B7.*

capítulo 14

MÉTODOS NO PARAMÉTRICOS

Objetivos

- Probar hipótesis cuando no es posible hacer alguna suposición sobre la distribución de la cual se muestrea
- Saber qué pruebas sin distribución (no paramétricas) son apropiadas para diferentes situaciones
- Usar e interpretar cada una de las seis pruebas de hipótesis no paramétricas estándar
- Conocer las ventajas y desventajas de las pruebas no paramétricas

Contenido del capítulo

14.1 Introducción a la estadística no paramétrica 622
14.2 Prueba de signo para datos por pares 624
14.3 Pruebas de suma de rangos: prueba *U* de Mann-Whitney y prueba de Kruskal-Wallis 630
14.4 Prueba de corridas de una sola muestra 640
14.5 Correlación de rango 646
14.6 Prueba de Kolmogorov-Smirnov 655

- Estadística en el trabajo 659
- Ejercicio de base de datos computacional 660
- Del libro de texto al mundo real 661
- Términos introducidos en el capítulo 14 662
- Ecuaciones introducidas en el capítulo 14 662
- Ejercicios de repaso 663

Aun cuando el efecto de la contaminación del aire sobre la salud es un problema complejo, una organización internacional ha decidido realizar una investigación preliminar de la calidad promedio del aire todo el año y de la incidencia de enfermedades pulmonares. Un estudio preliminar clasificó 11 de las principales ciudades del mundo desde 1 (peor) hasta 11 (mejor) para estas dos variables.

	Ciudad										
	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**
Clasificación de la calidad del aire	4	7	9	1	2	10	3	5	6	8	11
Clasificación de enfermedades pulmonares	5	4	7	3	1	11	2	10	8	6	9

Los datos de la organización de salud son distintos de los que hemos visto aquí: no nos proporcionan la variable usada para determinar estas clasificaciones o rangos. (No sabemos si la clasificación de enfermedades pulmonares es un resultado de neumonía, enfisema u otras afecciones para una población de 100,000.) Tampoco conocemos los valores (si la ciudad D tiene el doble de población que la ciudad K o 20 veces más). Si conociéramos las variables y sus valores, podríamos utilizar las técnicas de regresión del capítulo 12.

Desafortunadamente, no es así; pero aun sin conocer las variables o valores, podemos utilizar las técnicas de este capítulo para ayudar a la organización de salud con su problema.

14.1 Introducción a la estadística no paramétrica

Estadística paramétrica

Limitaciones de la estadística paramétrica

La mayoría de las pruebas de hipótesis analizadas hasta ahora hacen inferencias respecto a los *parámetros* de la población, como la media y la proporción. Estas pruebas paramétricas usan la estadística paramétrica de muestras que provinieron de la población que se está probando. Para formular estas pruebas, hicimos suposiciones restrictivas sobre las poblaciones de las que extraíamos nuestras muestras. En cada caso visto en los capítulos 8 y 9, por ejemplo, supusimos que nuestras muestras eran grandes o provenían de poblaciones con *distribución normal*. Pero las poblaciones no siempre son normales. Incluso cuando una prueba de bondad de ajuste (capítulo 11) indica que una población es aproximadamente normal, no siempre podemos estar seguros de que es correcto, porque la prueba no es 100% confiable. Claramente, existen ciertas situaciones en las que el uso de la curva normal no es apropiado. Para estos casos, requerimos alternativas a la estadística paramétrica y a las pruebas de hipótesis específicas que hemos estado utilizando hasta ahora.

Estadística no paramétrica

Por fortuna, recientemente los estadísticos han desarrollado técnicas útiles que no hacen suposiciones restrictivas respecto a la forma de las distribuciones de población. **Éstas se conocen como pruebas *sin distribución* o, más comúnmente, *pruebas no paramétricas.*** Las hipótesis de una prueba no paramétrica se refieren a algo distinto del valor de un parámetro de población. Existe un gran número de pruebas de este tipo, pero este capítulo examinará sólo algunas de las más conocidas y más utilizadas:

1. La *prueba de signo* para datos por pares, en la que los signos positivo o negativo sustituyen a los valores cuantitativos.
2. Una *prueba de suma de rangos*, a menudo llamada la *prueba U de Mann-Whitney*, que puede usarse para determinar si dos muestras independientes se sacaron de la misma población. Usa más información que la prueba de signo.

3. Otra prueba de suma de rangos, la *prueba Kruskal-Wallis*, que generaliza el análisis de varianza estudiado en el capítulo 11 para poder prescindir de la suposición de que las poblaciones tienen distribución normal.
4. La *prueba de corridas de una sola muestra*, un método para determinar la aleatoriedad con la que se han seleccionado los elementos muestreados.
5. *Correlación de rango*, un método para hacer el análisis de correlación cuando no se dispone de los datos para usar la forma numérica, pero cuando la información es suficiente para clasificar los datos como primero, segundo, tercero, etcétera.
6. La *prueba de Kolmogorov-Smirnov*, otro método para determinar la bondad de ajuste entre una muestra observada y una distribución de probabilidad teórica.

Ventajas de los métodos no paramétricos

Ventajas de los métodos no paramétricos

Los métodos no paramétricos tienen ciertas ventajas claras sobre los métodos paramétricos:

1. **No requieren la suposición de que una población está distribuida en forma de curva normal u otra forma específica.**
2. **Generalmente, es más sencillo realizarlas y entenderlas.** La mayor parte de las pruebas no paramétricas no exigen el tipo de cálculos laboriosos a menudo necesarios, por ejemplo, para calcular una desviación estándar. Una prueba no paramétrica nos puede pedir reemplazar valores numéricos con el orden en el que aparecen en una lista, como se ha hecho en la tabla 14-1. Obviamente, en términos de cálculos, manejar 1, 2, 3, 4 y 5 requiere menos esfuerzo que trabajar con 13.33, 76.50, 101.79, 113.45 y 189.42.
3. **Algunas veces ni siquiera se requiere un ordenamiento o clasificación formal.** Muchas veces, lo único que podemos hacer es describir un resultado como "mejor" que otro. Cuando ocurre esto, o cuando nuestras mediciones no son tan exactas como es necesario para las pruebas paramétricas, podemos usar métodos no paramétricos.

Desventajas de los métodos no paramétricos

Limitaciones de los métodos no paramétricos

Dos desventajas acompañan al uso de pruebas no paramétricas:

1. **Ignoran cierta cantidad de información.** Hemos mostrado cómo los valores 1, 2, 3, 4 y 5 pueden reemplazar a los números 13.33, 76.50, 101.79, 113.45 y 189.42. Sin embargo, si representamos "189.42" por "5" perdemos información contenida en el valor 189.42. Observe que en nuestro ordenamiento de los números 13.33, 76.50, 101.79, 113.45 y 189.42, el valor 189.42 puede convertirse en 1,189.42 y seguir siendo el quinto de la lista o el de mayor valor. Pero si esta lista es un conjunto de datos, tendremos un conocimiento mayor si sabemos que el valor más alto es 1,189.42, en vez de 189.42, de lo que sabríamos al representar ambos números con 5.
2. **A menudo no son tan eficientes o "claras" como las pruebas paramétricas.** La estimación de un intervalo a un nivel de confianza del 95% usando una prueba no paramétrica puede tener el doble de tamaño que la estimación con una prueba paramétrica como las de los capítulos 8 y 9. Cuando usamo pruebas no paramétricas hacemos un trueque: perdemos agudeza en la estimación de intervalos, pero ganamos la posibilidad de usar menos información y calcular con mayor rapidez.

Tabla 14-1
Conversión de valores paramétricos a rangos no paramétricos

Valor paramétrico	113.45	189.42	76.50	13.33	101.79
Valor no paramétrico	4	5	2	1	3

Ejercicios 14.1

Conceptos básicos

- **14-1** ¿Cuál es la diferencia entre los tipos de preguntas contestadas por pruebas paramétricas y las contestadas por pruebas no paramétricas?
- **14-2** La hipótesis nula examinada con mayor frecuencia en las pruebas no paramétricas (*elija un inciso*)
  a) incluye especificaciones de parámetros de una población.
  b) se usa para evaluar algún aspecto general de la población.
  c) es muy parecida a la utilizada en el análisis de regresión.
  d) prueba simultáneamente más que dos parámetros de población.
- **14-3** ¿Cuáles son las principales ventajas de los métodos no paramétricos sobre los métodos paramétricos?
- **14-4** ¿Cuáles son las principales desventajas de las pruebas no paramétricas?
- **14-5** George Shoaf es un entrevistador que trabaja en la oficina matriz de una aseguradora grande; para aprovechar mejor su tiempo, la compañía requiere que la recepcionista programe sus entrevistas de acuerdo con un programa preciso. No hay periodos de 5 minutos que no cuenten, ni siquiera de llamadas telefónicas. Desafortunadamente, la recepcionista minimizó la duración de las entrevistas, y ha programado demasiados solicitantes de empleo, lo que lleva a largas esperas en el vestíbulo. Aunque los periodos de espera pueden ser cortos por la mañana, al avanzar el día el entrevistador se va retrasando y se hacen más largos. Para evaluar el problema, ¿debe suponer el entrevistador que los tiempos de espera sucesivos siguen una distribución normal?

Aplicaciones

- **14-6** La International Communications Corporation planea cambiar el paquete de prestaciones ofrecido a los empleados. La compañía está considerando diferentes combinaciones de participación de utilidades, servicios médicos y planes de retiro. Algunas muestras de una amplia variedad de combinaciones de prestaciones se describieron en un folleto y se distribuyeron entre los empleados, luego se registraron a sus preferencias. Los siguientes fueron los resultados:

Rango	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Combinación de reparto de utilidades-Servicios médicos-retiro	15	5	14	4	6	16	7	8	13	3	17	18	12	2	9	1	11	19	10
Número de preferencias	52	49	39	38	37	36	32	29	26	25	24	18	15	15	14	10	10	10	9

¿Sacrificará la compañía información real al usar la prueba de rangos como su criterio de decisión? (*Sugerencia:* grafique los datos.)

14.2 Prueba de signo para datos por pares

Uso de la prueba de signo para datos apareados

Una de las pruebas no paramétricas más fáciles es la prueba de signo. Su nombre se debe a que está basada en la dirección (o signo de más o menos) de un par de observaciones y no en su magnitud numérica.

Considere el resultado de un panel de prueba de 40 estudiantes de penúltimo año de universidad que evalúa la efectividad de dos tipos de clases: conferencias grandes de profesores de tiempo com-

Tabla 14-2
Evaluación de dos tipos de clases por 40 estudiantes

Número del miembro de panel	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Calif. para conferencias grandes (1)	2	1	4	4	3	3	4	2	4	1	3	3	4	4	4	1
Calif. para secciones pequeñas (2)	3	2	2	3	4	2	2	1	3	1	2	3	4	4	3	2
Signo de la calificación:	−	−	+	+	−	+	+	+	+	0	+	0	0	0	+	−

pleto o secciones pequeñas con ayudantes de posgrado. La tabla 14-2 enumera las respuestas a esta petición: "Califique la efectividad de transmisión de conocimientos de estos dos tipos de clases, asignándo un número de 4 a 1. La calificación de 4 es excelente y el 1 es mala." En este caso, la prueba de signo nos puede ayudar a determinar si los estudiantes sienten que hay una diferencia entre la efectividad de los dos tipos de clases.

Conversión de valores a signos

Podemos comenzar, como en la tabla 14-2, convirtiendo las evaluaciones de los dos métodos de enseñanza en signos. Aquí, un signo más significa que el estudiante prefiere las conferencias grandes; un signo menos indica una preferencia por secciones pequeñas, y un cero representa un empate (sin preferencia). Si contamos el último renglón de la tabla 14-2, obtenemos lo siguiente:

Número de signos +	19
Número de signos −	11
Número de ceros	10
Tamaño total de muestra	**40**

Establecimiento de las hipótesis

Búsqueda del tamaño de muestra

Estamos usando la prueba de signo para determinar si nuestro panel percibe una diferencia real entre los dos tipos de clases. Puesto que estamos probando diferencias percibidas, excluiremos las evaluaciones empatadas (0). Vemos que hay 19 signos más y 11 signos menos, para un total de 30 respuestas utilizables. Si no hay diferencia entre los dos tipos de clases, p (la probabilidad de que la primera calificación exceda a la segunda) sería 0.5, y esperaríamos obtener aproximadamente 15 signos más y 15 signos menos. Estableceríamos nuestras hipótesis de la siguiente manera:

$H_0: p = 0.5$ ← Hipótesis nula: no hay diferencia entre los dos tipos de clases

$H_1: p \neq 0.5$ ← Hipótesis alternativa: hay diferencia entre los dos tipos de clases

Selección de la distribución

Si observa con cuidado las hipótesis, notará que la situación es parecida al lanzamiento de monedas que analizamos en el capítulo 4. Si lanzáramos al aire una moneda no cargada 30 veces, p sería 0.5, y esperaríamos aproximadamente 15 caras y 15 cruces. En ese caso, utilizaríamos la distribución binomial como la distribución de muestreo apropiada. Tal vez también recuerde que cuando np y nq son cada una al menos 5, podemos utilizar la distribución normal para aproximar a la binomial. Éste es precisamente el caso con los resultados de nuestro panel de estudiantes. Así, podemos aplicar la distribución normal a nuestra prueba de los dos métodos de enseñanza.

Planteamiento del problema en símbolos

$p_{H_0} = 0.5$ ← Proporción hipotética de la población que prefiere las conferencias para todos

$q_{H_0} = 0.5$ ← Proporción hipotética de la población que piensa que prefiere secciones pequeñas ($q_{H_0} = 1 - p_{H_0}$)

$n = 30$ ← Tamaño de muestra

$\bar{p} = 0.633$ ← Proporción de éxitos en la muestra (19/30)

$\bar{q} = 0.367$ ← Proporción de fracasos en la muestra (11/30)

Prueba de una hipótesis de que no hay diferencia

Cálculo del error estándar

Suponga que la oficina del rector desea probar la hipótesis de que no hay diferencia en la percepción del estudiante entre los dos tipos de clases al nivel de significancia de 0.05. Haremos esta

17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
1	2	2	4	4	4	4	3	3	2	3	4	3	4	3	1	4	3	2	2	2	1	3	3
3	2	3	3	1	4	3	3	2	2	1	1	1	3	2	2	4	4	3	3	1	1	4	2
−	0	−	+	+	0	+	0	+	0	+	+	+	+	+	−	0	−	−	−	+	0	−	+

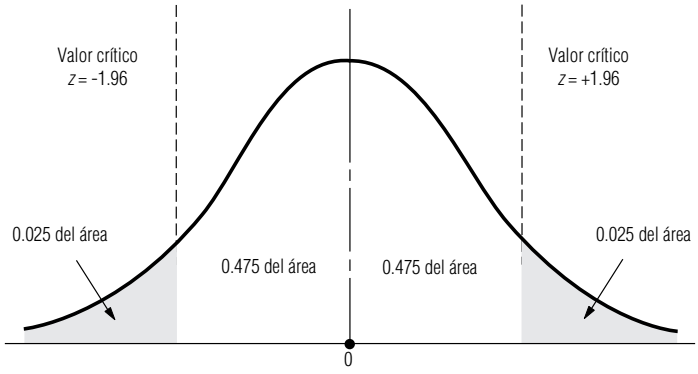


FIGURA 14-1

Prueba de hipótesis de dos colas de una proporción al nivel de significancia de 0.05

prueba usando los métodos del capítulo 8. El primer paso es calcular el error estándar de la proporción:

$$\sigma_{\bar{p}} = \sqrt{\frac{pq}{n}} \qquad [7\text{-}4]$$

$$= \sqrt{\frac{(0.5)(0.5)}{30}}$$

$$= \sqrt{0.00833}$$

$$= 0.091 \leftarrow \text{Error estándar de la proporción}$$

Representación gráfica de la prueba

Como queremos saber si la proporción verdadera es mayor o menor que la proporción hipotética, se trata de una prueba de dos colas. La figura 14-1 ilustra esta prueba de hipótesis. Las dos regiones sombreadas representan el nivel de significancia de 0.05.

Ahora usaremos la ecuación 6-2 para *estandarizar* la proporción de la muestra, $\bar{p}$, restando p_{H_0}, la proporción hipotética, y dividiendo entre $\sigma_{\bar{p}}$, el error estándar de la proporción.

$$z = \frac{\bar{p} - p_{H_0}}{\sigma_{\bar{p}}} \qquad [6\text{-}2]$$

$$= \frac{0.633 - 0.5}{0.091}$$

$$= 1.462$$

Interpretación de los resultados

Al colocar este valor estándar, 1.462, en la escala *z* se ve que la proporción de la muestra se encuentra dentro de la región de aceptación, como se ilustra en la figura 14-2. Por tanto, el rector debe aceptar la hipótesis nula de que los estudiantes no perciben diferencia entre los dos tipos de clase.

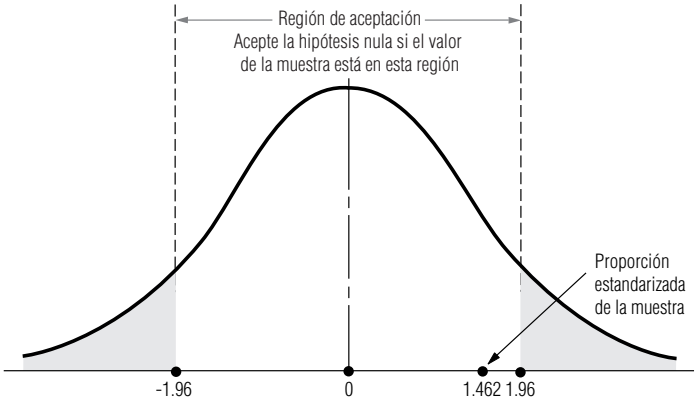


FIGURA 14-2

Prueba de hipótesis de dos colas al nivel de significancia de 0.05; ilustra la región de aceptación y la proporción estandarizada de la muestra

Una prueba de signo como ésta es bastante sencilla y se aplica para pruebas de una y de dos colas. Por lo general se basa en la distribución binomial. Sin embargo, recuerde que fue posible usar la aproximación normal a la binomial como distribución de la muestra porque *np* y *nq* eran mayores o iguales que 5. Cuando estas condiciones no se cumplen, debemos usar la distribución binomial.

SUGERENCIAS Y SUPOSICIONES

Las pruebas *no paramétricas* son muy convenientes cuando el mundo real presenta datos sin una distribución, con los cuales debe tomarse una decisión. Sugerencia: observe que la *prueba de signo* es sólo otra aplicación de la familiar *aproximación normal a la binomial*, que usa los signos + y − en lugar de "éxitos" y "fracasos".

Ejercicios 14.2

Ejercicios de autoevaluación

EA 14-1 Los siguientes datos muestran los índices de trabajo defectuoso de los empleados antes y después de un cambio en el plan de incentivos al salario. Compare los dos conjuntos de datos siguientes para ver si el cambio disminuyó las unidades defectuosas producidas. Utilice el nivel de significancia de 0.10.

Antes	8	7	6	9	7	10	8	6	5	8	10	8
Después	6	5	8	6	9	8	10	7	5	6	9	8

EA 14-2 Después de recolectar datos de la cantidad de contaminación del aire en Los Ángeles, la Agencia de Protección Ambiental (APA) decidió emitir nuevas reglas estrictas para controlar la cantidad de hidrocarburos en el aire. Para el año siguiente, tomó mediciones mensuales de este contaminante y las comparó con las mediciones del año anterior para los meses correspondientes. Con base en los siguientes datos, ¿tiene la APA suficiente evidencia para concluir con un 95% de confianza que las nuevas reglas fueron efectivas para disminuir la cantidad de hidrocarburos en el aire? Para justificar estas leyes para otro año, debe concluir, con $\alpha = 0.10$, que son efectivas. ¿Seguirán vigentes estas leyes el año siguiente?

	Año anterior*	Este año
Ene.	7.0	5.3
Feb.	6.0	6.1
Mar.	5.4	5.6
Abr.	5.9	5.7
May.	3.9	3.7
Jun.	5.7	4.7
Jul.	6.9	6.1
Ago.	7.6	7.2
Sep.	6.3	6.4
Oct.	5.8	5.7
Nov.	5.1	4.9
Dic.	5.9	5.8

*Medidos en partes por millón

Aplicaciones

■ **14-7** Los siguientes datos muestran los niveles de satisfacción de los empleados (como porcentaje) antes y después de que una empresa más grande compró a su compañía. ¿La compra aumentó la satisfacción de los empleados? Use un nivel de significancia de 0.05.

Antes	98.4	96.6	82.4	96.3	75.4	82.6	81.6	91.4	90.4	92.4
Después	82.4	95.4	94.2	97.3	77.5	82.5	81.6	84.5	89.4	90.6

■ **14-8** Use la prueba de los signos para ver si hay una diferencia entre el número de días requeridos para saldar una cuenta por cobrar antes y después de una nueva política de cobro. Use el nivel de significancia de 0.05.

Antes	33	36	41	32	39	47	34	29	32	34	40	42	33	36	29
Después	35	29	38	34	37	47	36	32	30	34	41	38	37	35	28

■ **14-9** Un taller de reparación para motores de avionetas cambió el método de pago de salario por hora a salario por hora más un bono calculado sobre el tiempo requerido para desmontar, reparar y volver a ensamblar un motor. Los siguientes son datos recabados para 25 motores antes del cambio y 25 después. A un nivel de significancia de 0.10, ¿el nuevo plan incrementó la productividad?

Horas requeridas		Horas requeridas	
Antes	**Después**	**Antes**	**Después**
29	32	25	34
34	19	42	27
32	22	20	26
19	21	25	25
31	20	33	31
22	24	34	19
28	25	20	22
31	31	21	32
32	18	22	31
44	22	45	30
41	24	43	29
23	26	31	20
34	41		

■ **14-10** Debido a la severidad de los inviernos recientes, se ha dicho que la Tierra progresa lentamente hacia otra era glacial. Sin embargo, algunos científicos tienen otros puntos de vista, porque los veranos también han traído temperaturas extremas. Un científico sugirió examinar la temperatura media de cada mes para ver si era menor que la del año anterior. Otro miembro del servicio meteorológico gubernamental argumentó que tal vez debían revisar también las temperaturas de los meses de primavera y otoño por lo menos de los últimos dos años, de forma que las conclusiones no sólo se basaran en las temperaturas extremas. De esta manera, dijo, podrían detectar la posibilidad de una tendencia general de calentamiento o enfriamiento o sólo temperaturas extremas en los meses de verano e invierno. Se seleccionaron al azar 15 fechas de primavera y otoño, y se anotaron las temperaturas de los últimos dos años para un lugar particular con temperaturas generalmente moderadas. A continuación se dan las fechas y temperaturas correspondientes a 1994 y 1995.

a) ¿Es correcto el razonamiento del meteorólogo como método de evaluación? Explique su respuesta.
b) Usando la prueba de los signos, determine si el meteorólogo puede concluir, con $\alpha = 0.05$, que 1995 fue más frío que 1994 con base en estos datos.

	Temperatura (Fahrenheit)			Temperatura (Fahrenheit)	
Fecha	**1994**	**1995**	**Fecha**	**1994**	**1995**
Mar. 29	58	57	Oct. 12	54	48
Abr. 4	45	70	May 31	74	79
Abr. 13	56	46	Sept. 28	69	60
May. 22	75	67	Jun. 5	80	74
Oct. 1	52	60	Jun. 17	82	79
Mar. 23	49	47	Oct. 5	59	72
Nov. 12	48	45	Nov. 28	50	50
Sept. 30	67	71			

■ **14-11** Preocupados por la exposición a la radiación y su relación con la incidencia de cáncer, los especialistas ambientales urbanos mantienen una estrecha vigilancia sobre el tipo de industria que llega al área y el grado en el que usan radiación en su producción. Se ha desarrollado un índice de exposición a la contaminación radiactiva y se utiliza todos los días para determinar si los niveles están aumentando o son más altos para ciertas condiciones atmosféricas.

Los ambientalistas afirman que la contaminación radiactiva se ha incrementado en el último año debido a las nuevas industrias en la ciudad. Sin embargo, los administradores de la ciudad afirman que el nuevo reglamento más estricto sobre la industria en el área ha bajado los niveles respecto al año anterior, aun con la nueva industria que usa radiación. Para probar su afirmación, se verificaron registros de 11 días aleatorios y se anotó el índice de exposición a la contaminación radiactiva. Se obtuvieron los siguientes resultados:

	Índice de exposición a la radiación										
1994	1.402	1.401	1.400	1.404	1.395	1.402	1.406	1.401	1.404	1.406	1.397
1995	1.440	1.395	1.398	1.404	1.393	1.400	1.401	1.402	1.400	1.403	1.402

¿Pueden los administradores concluir, para $\alpha = 0.15$, que los niveles de contaminación radiactiva han cambiado o, más específicamente, que se han reducido?

■ **14-12** Como parte del reciente interés en el crecimiento de la población y el tamaño de las familias, una investigadora demográfica examinó cierto número de hipótesis referentes al tamaño de familia que diversas personas consideran como ideal. Sospechaba que las variables de raza, edad y educación podrían explicar los distintos puntos de vista. En una muestra piloto, la investigadora probó la hipótesis de que las mujeres actuales piensan que la familia ideal debe ser de tamaño menor que la que sus madres consideraban ideal. Le pidió a cada participante del estudio piloto que estableciera el número de hijos que elegiría tener o que consideraba ideal. Las respuestas fueron anónimas, para evitar la posibilidad de que la gente se sintiera obligada a dar una respuesta socialmente deseable. Además, la muestra incluyó personas con distintos antecedentes. Las siguientes son las respuestas de las parejas madre-hija.

	Tamaño ideal de familia												
Pareja muestra	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**
Hija	3	4	2	1	5	4	2	2	3	3	1	4	2
Madre	4	4	4	3	5	3	3	5	3	2	2	3	1

a) ¿Puede la investigadora concluir, con $\alpha = 0.03$, que las madres e hijas no tienen esencialmente el mismo ideal de tamaño de familia? Use la distribución binomial.
b) Determine si la investigadora podría llegar a la conclusión de que las madres no tienen esencialmente las mismas preferencias de tamaño de familia que sus hijas usando la aproximación normal a la binomial.
c) Suponga que por cada pareja enumerada, hubo otras 10 parejas que respondieron de manera idéntica. Calcule el intervalo de la proporción para el cual la investigadora puede concluir que no hay diferencia entre madres e hijas. ¿Es distinta su conclusión?
d) Explique cualquier diferencia en las conclusiones de los incisos a), b) y c).

■ **14-13** Una compañía que vende automóviles usados en todo el país desarrolló un nuevo video para capacitar a su personal de ventas. Los datos presentan los promedios mensuales de ventas de 20 empleados para periodos tanto anteriores como posteriores a la creación del video. ¿Tiene la compañía suficiente evidencia para concluir, con el 95% de confianza, que el video fue efectivo para aumentar el número promedio de autos vendidos? Si sólo se considera a los empleados con bajas ventas (menos que un promedio de 12 autos por mes antes del video), ¿el video aumentó su desempeño de ventas?

Antes	18.4	16.9	17.4	11.6	10.5	12.7	22.3	18.5	17.5	16.4
Después	18.6	16.8	17.3	15.6	19.5	12.6	22.3	16.5	18.0	16.4
Antes	15.9	18.6	23.5	18.7	9.4	16.3	18.5	17.4	11.3	8.4
Después	17.4	18.6	23.5	18.9	15.6	15.4	17.6	17.4	16.5	13.4

Soluciones a los ejercicios de autoevaluación

EA 14-1

Antes	8	7	6	9	7	10	8	6	5	8	10	8
Después	6	5	8	6	9	8	10	7	5	6	9	8
Signo	−	−	+	−	+	−	+	+	0	−	−	0

12 respuestas: 4(+), 6(−), 2(0).

Para $n = 10$, $p = 0.5$, la probabilidad de 6 o más signos menos es 0.3770 (tabla 3 del apéndice). Como $0.3770 > 0.10$, no se puede rechazar H_0. El plan de incentivos al salario no baja significativamente las tasas de trabajo defectuoso.

EA 14-2

Antes	7.0	6.0	5.4	5.9	3.9	5.7	6.9	7.6	6.3	5.8	5.1	5.9
Después	5.3	6.1	5.6	5.7	3.7	4.7	6.1	7.2	6.4	5.7	4.9	5.8
Signo	−	+	+	−	−	−	−	−	+	−	−	−

12 respuestas: 3(+), 9(−).

Para $n = 12$, $p = 0.5$, la probabilidad de 9 o más signos menos es 0.0729 (tabla 3 del apéndice). Como $0.10 > 0.0729 > 0.05$, no pueden tener una confianza del 95% de que los niveles de hidrocarburos hayan bajado, pero pueden concluir, para $\alpha = 0.10$, que las reglas son efectivas. Así, estarán vigentes todavía el año próximo.

14.3 Pruebas de suma de rangos: prueba *U* de Mann-Whitney y prueba de Kruskal-Wallis

En el capítulo 11, mostramos cómo usar el análisis de varianza para probar la hipótesis de que varias medias de población son iguales. Supusimos para esas pruebas que las poblaciones tenían una distribución normal con varianzas iguales. Muchas veces estas suposiciones no se satisfacen y entonces podemos utilizar dos pruebas no paramétricas; ninguna de ellas depende de las suposiciones de normalidad. Estas dos pruebas se llaman pruebas de suma de rangos porque la prueba depende de los rangos o clasificaciones de las observaciones de muestra.

Se usan según el número de poblaciones involucradas

Las pruebas de suma de rangos son una familia completa de pruebas; nos concentraremos en sólo dos miembros de ella: la prueba *U* de Mann-Whitney y la prueba de Kruskal-Wallis. Usaremos la prueba de Mann-Whitney cuando tengamos sólo dos poblaciones y la prueba de Kruskal-Wallis cuando se trate de más de dos. El uso de estas pruebas nos permitirá determinar si las muestras independientes se obtuvieron de la misma población (o de distintas poblaciones con la misma distribución). El uso de la *clasificación* de la información en lugar de los signos más y menos desperdicia menos datos que la prueba de los signos.

Solución de un problema usando la prueba *U* de Mann-Whitney

Suponga que la junta directiva de una gran universidad estatal del este de Estados Unidos desea probar la hipótesis de que las calificaciones promedio en la prueba SAT en dos planteles de la universidad son iguales. La junta mantiene estadísticas de todos los estudiantes en todos los planteles del sistema. Una muestra aleatoria de 15 estudiantes de cada plantel produjo los datos que se muestran en la tabla 14-3.

Clasificación de los elementos que se van a probar

Para aplicar la prueba *U* de Mann-Whitney a este problema, comenzamos por clasificar las calificaciones en orden ascendente, indicando junto a cada una el símbolo del plantel, como se ve en la tabla 14-4.

A continuación, aprendamos los símbolos usados para la prueba *U* de Mann-Whitney en el contexto de este problema:

Símbolos para expresar el problema

n_1 = número de elementos en la muestra 1, es decir, el número de estudiantes del plantel A

n_2 = número de elementos en la muestra 2, es decir, el número de estudiantes del plantel S

R_1 = suma de los rangos de los elementos en la muestra 1: la suma de los rangos de todas las calificaciones del plantel A de la tabla 14-5

R_2 = suma de los rangos de los elementos en la muestra 2: la suma de los rangos de todas las calificaciones del plantel S de la tabla 14-5

En este caso, tanto n_1 como n_2 son iguales a 15, pero *no* es necesario que ambas muestras sean del mismo tamaño. Ahora, en la tabla 14-5, reproducimos los datos de la tabla 14-3, agregando los

Tabla 14-3
Calificaciones en el SAT de estudiantes en dos planteles de la universidad estatal

Plantel A	1,000	1,100	800	750	1,300	950	1,050	1,250
Plantel S	920	1,120	830	1,360	650	725	890	1,600
Plantel A	1,400	850	1,150	1,200	1,500	600	775	
Plantel S	900	1,140	1,550	550	1,240	925	500	

Tabla 14-4
Calificaciones en el SAT clasificadas de menor a mayor

Rango	Calif.	Plantel	Rango	Calif.	Plantel
1	500	S	16	1,000	A
2	550	S	17	1,050	A
3	600	A	18	1,100	A
4	650	S	19	1,120	S
5	725	S	20	1,140	S
6	750	A	21	1,150	A
7	775	A	22	1,200	A
8	800	A	23	1,240	S
9	830	S	24	1,250	A
10	850	A	25	1,300	A
11	890	S	26	1,360	S
12	900	S	27	1,400	A
13	920	S	28	1,500	A
14	925	S	29	1,550	5
15	950	A	30	1,600	S

rangos de la tabla 14-4. Después sumamos los rangos de cada plantel. Como resultado, tenemos todos los valores que necesitamos para solucionar este problema, porque sabemos que

$$n_1 = 15$$
$$n_2 = 15$$
$$R_1 = 247$$
$$R_2 = 218$$

Cálculo del estadístico *U*

Definición del estadístico *U*

Usando los valores de n_1 y n_2 y las sumas de rangos R_1 y R_2, podemos determinar el estadístico U, una medida de la diferencia entre las observaciones clasificadas de las dos muestras de calificaciones en el SAT:

Cálculo del estadístico *U*

Estadístico *U*

$$U = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1 \qquad [14\text{-}1]$$

$$= (15)(15) + \frac{(15)(16)}{2} - 247$$
$$= 225 + 120 - 247$$
$$= 98 \leftarrow \text{Estadístico } U$$

Si la hipótesis nula de que $n_1 + n_2$ observaciones provienen de poblaciones idénticas es cierta, entonces el estadístico U tiene una distribución muestral con una media de

Tabla 14-5 Datos sin procesar y rango de las calificaciones en el SAT

Plantel A	Rango	Plantel S	Rango
1,000	16	920	13
1,100	18	1,120	19
800	8	830	9
750	6	1,360	26
1,300	25	650	4
950	15	725	5
1,050	17	890	11
1,250	24	1,600	30
1,400	27	900	12
850	l0	1,140	20
1,150	21	1,550	29
1,200	22	550	2
1,500	28	1,240	23
600	3	925	14
775	7	500	1
	247 ← **Total de rangos**		**218** ← **Total de rangos**

Media de la distribución muestral de *U*

$$\mu_U = \frac{n_1 n_2}{2} \qquad [14\text{-}2]$$

$$= \frac{(15)(15)}{2}$$

$$= 112.5 \leftarrow \text{Media del estadístico } U$$

y un error estándar de

Error estándar del estadístico *U*

$$\sigma_U = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}} \qquad [14\text{-}3]$$

$$= \sqrt{\frac{(15)(15)(15 + 15 + 1)}{12}}$$

$$= \sqrt{\frac{6{,}975}{12}}$$

$$= \sqrt{581.25}$$

$$= 24.1 \leftarrow \text{Error estándar del estadístico } U$$

Prueba de las hipótesis

La distribución muestral del estadístico U puede aproximarse por la distribución normal cuando tanto n_1 como n_2 son mayores que 10. Como nuestro problema cumple esta condición, podemos usar la tabla de distribución de probabilidad normal estándar para hacer la prueba. La junta de directores de-

sea probar al nivel de significancia de 0.15 la hipótesis de que estas muestras fueron extraídas de poblaciones idénticas.

Planteamiento de las hipótesis

H_0: $\mu_1 = \mu_2$ ← Hipótesis nula: no hay diferencia entre las dos poblaciones, tienen la misma media

H_1: $\mu_1 \neq \mu_2$ ← Hipótesis alternativa: existe una diferencia entre las dos poblaciones; en particular, tienen distintas medias

$\alpha = 0.15$ ← Nivel de significancia para probar estas hipótesis

Búsqueda de los límites de la región de aceptación

La junta de directores desea saber si el promedio de calificación de la prueba de aptitud académica, SAT, en cualquiera de dos escuelas es mejor o peor que en la otra. La figura 14-13 ilustra la gráfica de esta prueba. Las dos áreas sombreadas representan el nivel de significancia de 0.15. Como se está usando la distribución normal como distribución de la muestra en esta prueba, de la tabla 1 del apéndice se puede determinar que el valor crítico de *z* para un área de 0.425 es 1.44.

Ahora usamos la ecuación 6-2 para *estandarizar* la estadística U de la muestra, restando μ_U, su media, y dividiendo entre μ_U, su error estándar.

$$z = \frac{U - \mu_U}{\sigma_U} \qquad [6\text{-}2]$$

$$= \frac{98 - 112.5}{24.1}$$

$$= -0.602$$

La figura 14-4 señala el valor estandarizado de la muestra U y los valores críticos de *z* para la prueba. La junta de directores debe observar que el estadístico de la muestra está dentro de los valores críticos para la prueba y concluir que las distribuciones y, por ende, las medias de las calificaciones del SAT en las dos escuelas son iguales.

Propiedades especiales de la prueba *U*

Otra forma de calcular el estadístico *U*

El estadístico U tiene una característica que permite a los usuarios ahorrar tiempo en cálculos cuando las dos muestras bajo observación tienen diferente tamaño. Acabamos de calcular el valor de U usando la ecuación 14-1:

$$U = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1 \qquad [14\text{-}1]$$

Pero con la misma facilidad podríamos haber calculado la estadística U usando el valor de R_2, así:

Fórmula alternativa para el estadístico *U*

$$U = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2 \qquad [14\text{-}4]$$

FIGURA 14-3

Prueba de hipótesis de dos colas al nivel de significancia de 0.15

Representación gráfica de la prueba

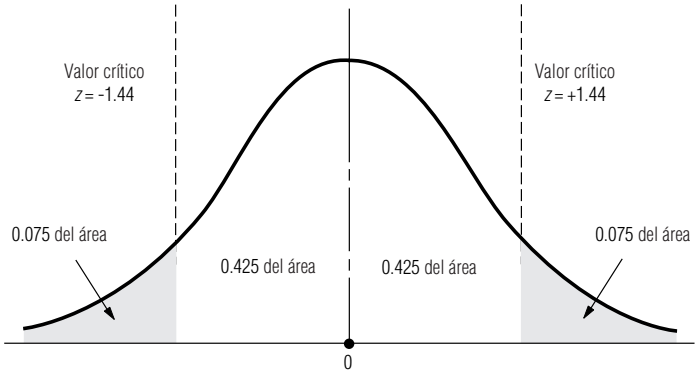
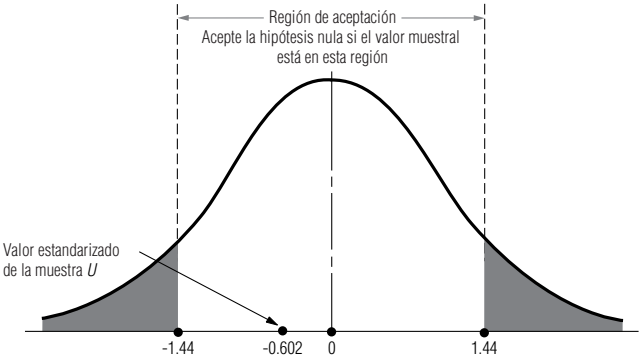


FIGURA 14-4

Prueba de hipótesis de dos colas al nivel de significancia de 0.15; señala la región de aceptación y el estadístico de la muestra *U*



La respuesta habría sido 127 (que se aleja hacia *arriba* de la media de 112.5 justo tanto como 98 se aleja hacia *abajo*). En este problema, habríamos tardado lo mismo calculando el valor del estadístico *U* con la ecuación 14-1 que con la 14-4. En otros casos, cuando el número de elementos es mayor en una muestra que en la otra, elija la ecuación que requerirá menos trabajo. Sin importar si calculó *U* usando la ecuación 14-1 o la 14-4, llegará a la misma conclusión. Observe que en este ejemplo, la respuesta 127 cae en la región de aceptación tal como lo hizo 98.

Manejo de empates en los datos

¿Qué pasa si aparecen *empates* cuando clasificamos los elementos para esta prueba? Por ejemplo, ¿qué sucede si las calificaciones 13 y 14 de la tabla 14-4 tuvieran ambas el valor 920? En este caso, encontramos el promedio de sus rangos (13 + l4)/2 = 13.5, y asignamos el resultado a los dos. Si hubiera un empate entre tres clasificaciones, 13, 14 y 15, por ejemplo, promediamos estos rangos (13 + 14 + 15)/3 = 14, y usamos ese valor para los tres elementos.

Solución de un problema usando la prueba de Kruskal-Wallis

Prueba de diferencias cuando están involucrados más de dos poblaciones

Como se observó en esta sección, la prueba de Kruskal-Wallis es una extensión de la prueba de Mann-Whitney para casos en que están involucradas más de dos poblaciones. Esta prueba, también, depende de los rangos de las observaciones de la muestra.

La tabla 14-6 presenta las calificaciones de una muestra de 20 pilotos estudiantes en su examen escrito de la Agencia Federal de Aviación (AFA), dispuestas según el método que se empleó en su capacitación: videocasete, audiocasete o salón de clase.

La AFA está interesada en evaluar la efectividad de estos tres métodos de capacitación. Específicamente, desea probar, al nivel de significancia de 0.10, la hipótesis de que las calificaciones medias en el examen escrito de los pilotos estudiantes capacitados por estos tres métodos son iguales. Puesto que tenemos más de dos poblaciones involucradas, es pertinente aplicar la prueba de Kruskal-Wallis en este caso. Para ello, comenzamos en la tabla 14-7 por clasificar todas las calificaciones en orden ascendente, indicando junto a cada una el símbolo del método de entrenamiento que se utilizó. Los empates se manejan promediando los rangos, como se hizo con la prueba de Mann-Whitney.

Clasificación de los elementos de la prueba

A continuación, definimos los símbolos usados en una prueba de Kruskal-Wallis:

Símbolos utilizados para una prueba de Kruskal-Wallis

n_j = número de elementos en la muestra j

R_j = suma de los rangos de todos los elementos en la muestra j

k = número de muestras

$n = n_1 + n_2 + \cdots + n_k$, el número total de observaciones en todas las muestras

Tabla 14-6

Calificaciones en el examen escrito de 20 pilotos estudiantes capacitados con tres métodos

Videocasete	74	88	82	93	55	70			
Audiocasete	78	80	65	57	89				
Salón	68	83	50	91	84	77	94	81	92

Tabla 14-7
Calificaciones del examen escrito clasificadas en orden ascendente

Rango	Calificaciones	Método de capacitación	Rango	Calificaciones	Método de capacitación
1	50	S	11	81	S
2	55	VC	12	82	VC
3	57	AC	13	83	S
4	65	AC	14	84	S
5	68	S	15	88	VC
6	70	VC	16	89	AC
7	74	VC	17	91	S
8	77	S	18	92	S
9	78	AC	19	93	VC
l0	80	AC	20	94	S

Reclasificación de los datos para calcular sumas de rangos

La tabla 14-8 contiene los datos de la tabla 14-7 reclasificados de tal forma que facilita calcular las sumas de los rangos para cada método de capacitación. Entonces podemos usar la ecuación 14-5 para calcular el estadístico *K*, una medida de las diferencias entre las observaciones clasificadas en las tres muestras.

Cálculo del estadístico *K*

Estadístico *K*

$$K = \frac{12}{n(n+1)} \sum_j \frac{R_j^2}{n_j} - 3(n+1) \qquad [14\text{-}5]$$

$$= \frac{12}{20(20+1)} \left[\frac{(61)^2}{6} + \frac{(42)^2}{5} + \frac{(107)^2}{9} \right] - 3(20+1)$$

$$= (0.02857)(620.2 + 352.8 + 1{,}272.1 - 63)$$

$$= 1.143$$

Prueba de las hipótesis

La distribución muestral del estadístico *K* puede aproximarse por una distribución ji-cuadrada ***cuando los tamaños de todas las muestras son al menos 5***. Como nuestro problema satisface esta condición, podemos usar la distribución ji-cuadrada y la tabla 5 del apéndice para esta prueba. En una

Tabla 14-8
Datos y rango ordenados según el método de capacitación

Videocasete	Rango	Audiocasete	Rango	Salón	Rango
74	7	78	9	68	5
88	15	80	10	83	13
82	12	65	4	50	1
93	19	57	3	91	17
55	2	89	16	84	14
70	6		**42** ← Suma	77	8
	61 ← Suma		de rangos	94	20
	de rangos			81	11
				92	18
					107 ← Suma de rangos

FIGURA 14-5

Prueba de Kruskal-Wallis al nivel de significancia de 0.10; señala la región de aceptación y el estadístico muestral *K*

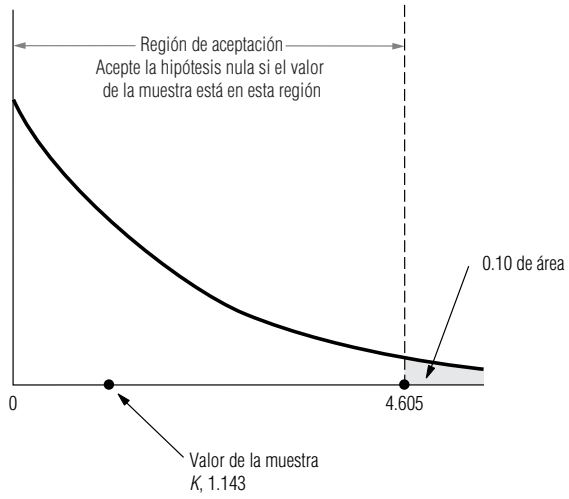


Ilustración de la prueba

prueba de Kruskal-Wallis, el número apropiado de grados de libertad es $k - 1$, que en este problema es $(3 - 1)$ o 2, puesto que tenemos tres muestras. Las hipótesis pueden enunciarse de la siguiente manera:

Formulación de las hipótesis

H_0: $\mu_1 = \mu_2 = \mu_3$ ← Hipótesis nula: no hay diferencia entre las tres poblaciones, es decir, tienen la misma media

H_1: μ_1, μ_2 y μ_3 no todas son iguales ← Hipótesis alternativa: existen diferencias entre las tres poblaciones; en particular, tienen distintas medias

$\alpha = 0.10$ ← Nivel de significancia para probar estas hipótesis

Interpretación de los resultados

La figura 14-5 ilustra una distribución ji-cuadrada con 2 grados de libertad. El área sombreada representa el nivel de significancia de 0.10. Observe que la región de aceptación para la hipótesis nula (que no existen diferencias entre las tres poblaciones) se extiende de cero a un valor ji-cuadrada de 4.605. Obviamente, el valor muestral *K* de 1.143 está dentro de la región de aceptación; por consiguiente, la AFA debe aceptar la hipótesis nula y concluir que no hay diferencias en los resultados obtenidos al usar los tres métodos de capacitación.

SUGERENCIAS Y SUPOSICIONES

Las pruebas de suma de rangos, como las de Mann-Whitney y de Kruskal-Wallis, a menudo producen empates. Cuando los encuentre, recuerde que cada valor empatado obtiene un rango *promedio*. Si los elementos 10° y 11° están empatados, cada uno obtiene un rango de 10.5. En el caso de empates de más de 2 elementos, todos obtienen el rango promedio [un empate en los elementos 3°, 4°, 5° y 6° significa que los cuatro obtienen el rango de $(3 + 4 + 5 + 6)/4 = 4.5$].

Ejercicios 14.3

Ejercicios de autoevaluación

EA 14-3 La boutique de Melisa tiene tres establecimientos en centros comerciales. Melisa mantiene un registro diario del número de clientes que realmente compran en cada establecimiento. La siguiente es una muestra de esos datos. Utilizando la prueba de Kruskal-Wallis, ¿puede decir, al nivel de significancia de 0.05, que sus tiendas tienen el mismo número de clientes que compran?

Centro comercial Eastowne	99	64	101	85	79	88	97	95	90	100
Centro comercial Craborchard	83	102	125	61	91	96	94	89	93	75
Centro comercial Fairforest	89	98	56	105	87	90	87	101	76	89

EA 14-4 Un gran hospital contrata la mayor parte de sus enfermeras en dos universidades importantes del área. Durante el año pasado, aplicaron un examen a las enfermeras recién graduadas que entran al hospital para

determinar qué escuela parece preparar mejor a sus enfermeras, si alguna lo hace. Basándose en las siguientes calificaciones (de 100 puntos posibles), ayude a la oficina de personal del hospital a determinar si las escuelas difieren en calidad. Use la prueba u de Mann-Whitney con un nivel de significancia del 10%.

	Calificaciones del examen												
Escuela A	97	69	73	84	76	92	90	88	84	87	93		
Escuela B	88	99	65	69	97	84	85	89	91	90	87	91	72

Aplicaciones

■ **14-14** Pruebe la hipótesis de que no hay diferencia entre las edades de empleados masculinos y femeninos de cierta compañía, utilizando la prueba u de Mann-Whitney para los datos de la muestra. Use el nivel de significancia de 0.10.

Hombres	31	25	38	33	42	40	44	26	43	35
Mujeres	44	30	34	47	35	32	35	47	48	34

■ **14-15** En la tabla siguiente se dan los precios de venta de tres marcas de zapatos. Use la prueba de Kruskal-Wallis para determinar si existe alguna diferencia entre los precios de venta de las marcas a lo largo del país. Utilice el nivel de significancia de 0.01.

Marca A	$89	90	92	81	76	88	85	95	97	86	100
Marca B	$78	93	81	87	89	71	90	96	82	85	
Marca C	$80	88	86	85	79	80	84	85	90	92	

■ **14-16** Una compañía de regalos de compras por catálogo tiene la siguiente muestra de ventas en dólares, separados de acuerdo con la forma de pago del pedido. Pruebe la hipótesis de que no existe diferencia en la cantidad en dólares recibida por pagos en efectivo, cheque o tarjeta de crédito. Use la prueba de Kruskal-Wallis con un nivel de significancia de 0.05.

Pagos con tarjeta de crédito	78	64	75	45	82	69	60
Pagos con cheque	110	70	53	51	61	68	
Pagos con efectivo	90	68	70	54	74	65	59

■ **14-17** Los siguientes datos muestran horas anuales perdidas por enfermedad de los 24 hombres y mujeres de la Northern Packing Company, Inc. Al nivel de significancia de 0.10, ¿existe alguna diferencia atribuible al sexo? Utilice la prueba *U* de Mann-Whitney.

Hombres	31	44	25	30	70	63	54	42	36	22	25	50
Mujeres	38	34	33	47	58	83	18	36	41	37	24	48

■ **14-18** Un fabricante de juguetes cambió el tipo de máquinas inyectoras de plástico que usaba porque una máquina moderna dio evidencias de ser más económica. Sin embargo, al iniciar la temporada de Navidad, la productividad pareció un poco menor a la del año anterior. Los registros de producción de los últimos años estaban disponibles y el gerente de producción decidió comparar el resultado mensual de los 15 meses en que se usaron las máquinas antiguas y los 11 meses de producción de ese año. Los registros muestran estas cantidades de producción con las máquinas antiguas y las nuevas.

Producción mensual en unidades		
Máquinas antiguas		**Máquinas nuevas**
992	966	965
945	889	1,054
938	972	912
1,027	940	850
892	873	796
983	1,016	911
1,014	897	877
1,258		902
		956
		900
		938

¿Puede concluir la compañía, a un nivel de significancia de 0.10, que el cambio en las máquinas ha reducido la producción?

■ **14-19** Hank's Hot Dogs tiene cuatro puestos de hot dogs en el Memorial Stadium. Hank sabe cuántos hot dogs se venden en cada puesto durante cada juego de fútbol y quiere determinar si los cuatro venden el mismo número. Use la prueba de Kruskal-Wallis, para un nivel de significancia de 0.10, ¿puede decir si los puestos tienen la misma cantidad de ventas de hot dogs?

Juego	1	2	3	4	5	6	7	8	9
Visitantes norte	755	698	725	895	886	794	694	827	814
Visitantes sur	782	724	754	825	815	826	752	784	789
Casa norte	714	758	684	816	856	884	774	812	734
Casa sur	776	824	654	779	898	687	716	889	917

■ **14-20** Para incrementar las ventas durante los días de más demanda, una cadena de tiendas que vende queso en centros comerciales da pruebas de su producto a la entrada de las tiendas. La administración de la cadena define los días de más venta y selecciona aleatoriamente los días para repartir muestras. De una muestra de días que se consideraron fuertes en compras, los siguientes datos proporcionan las ventas de una tienda en los días que dio muestras de quesos y en los que no.

Ventas (en cientos)													
Días de promoción	18	21	23	15	19	26	17	18	22	20	18	21	27
Días normales	22	17	15	23	25	20	26	24	16	17	23	21	

Use la prueba u de Mann-Whitney y un nivel de significancia de 5% para decidir si regalar muestras frente a las tiendas produjo mayores ventas.

■ **14-21** Una compañía está interesada en saber si hay una diferencia en la tasa de producción de empleados hombres y mujeres en el departamento de moldes. Se pidió a Judy Johnson, gerente de producción, que realizara un estudio para medir la producción de trabajadores hombres y mujeres durante una semana. De alguna manera, uno de los empleados de oficina extravió una porción de los datos, y Judy sólo pudo localizar la siguiente información de los registros de las pruebas:

$$\sigma_U = 176.4275$$

$$\mu_U = 1{,}624$$

$$R_1 = 3{,}255$$

Judy también recordó que el tamaño de la muestra de hombres, n_2, era 2 unidades mayor que n_1.

Reconstruya un valor z para la prueba y determine si puede suponerse que la producción semanal, a un nivel de significancia del 5%, es la misma tanto para hombres como para mujeres. También indique los valores para n_1, n_2 y R_2.

■ **14-22** Una universidad que acepta estudiantes de preparatorias rurales y urbanas desea saber si los antecedentes distintos llevan a una diferencia en los promedios generales del primer año. Se presentan los datos de 13 estudiantes de zona rural de primer ingreso elegidos al azar y de 16 con antecedentes urbanos. Use la prueba U de Mann-Whitney con el 5% de nivel de significancia.

Promedios generales del primer año										
Rural	3.19	2.05	2.82	2.16	3.84	4.0	2.91	2.75	3.01	1.98
	2.58	2.76	2.94							
Urbana	3.45	3.16	2.84	2.09	2.11	3.08	3.97	3.85	3.72	2.73
	2.81	2.64	1.57	1.87	2.54	2.62				

■ **14-23** Veinte vendedores de la compañía Henley Paper recibieron capacitación en ventas durante el año pasado. Algunos fueron enviados a un programa nacional conducido por maestros en ventas. Los otros se capacitaron en las oficinas de la compañía con el gerente de ventas de Henley. Se dan los porcentajes de las cuotas de ventas logrados por ambos grupos durante el último año. El señor Boyden Henley, presidente, cree que los antecedentes, las aptitudes de ventas y la motivación de ambos grupos son comparables. Al nivel de significancia de 0.10, ¿ha resultado mejor alguno de los dos métodos de capacitación? Use la prueba u de Mann-Whitney.

	Porcentaje de las cuotas logrado									
Maestros en ventas	90	95	105	110	100	75	80	90	105	120
Compañía	80	90	100	120	95	95	90	100	95	105

Soluciones a los ejercicios de autoevaluación

EA 14-3

Centro comercial Eastowne	99	64	101	85	79	88	97	95	90	100
	24	3	26.5	8	6	11	22	20	15.5	25
Centro comercial Craborchard	83	102	125	61	91	96	94	89	93	75
	7	28	30	2	17	21	19	13	18	4
Centro comercial Fairforest	89	98	56	105	87	90	87	101	76	89
	13	23	1	29	9.5	15.5	9.5	26.5	5	13

$n_1 = 10$ $n_2 = 10$ $n_3 = 10$ $\alpha = 0.05$
$R_1 = 161$ $R_2 = 159$ $R_3 = 145$
H_0: $\mu_1 = \mu_2 = \mu_3$ H_1: las μ no son todas iguales

$$K = \frac{12}{n(n+1)} \sum \frac{R_j^2}{n_j} - 3(n+1)$$

$$\frac{1R_j^2}{n\,n_j} = \frac{12}{30(31)}\left(\frac{(161)^2}{10} + \frac{(159)^2}{10} + \frac{(145)^2}{10}\right) - 3(31) = 0.196$$

Con $3 - 1 = 2$ grados de libertad y $\alpha = 0.05$, el límite superior de la región de aceptación es $\chi^2 = 5.991$, de manera que se acepta H_0. Los números de compradores promedio en los tres centros no son significativamente diferentes.

EA 14-4

Escuela A	97	69	73	84	76	92	90	88	84	87	93		
	22.5	2.5	5	8	6	20	16.5	13.5	8	11.5	21		
Escuela B	88	99	65	69	97	84	85	89	91	90	87	91	72
	13.5	24	1	2.5	22.5	8	10	15	18.5	16.5	11.5	18.5	4

$n_1 = 11$ $n_2 = 13$ $\alpha = 0.10$
$R_1 = 134.5$ $R_2 = 165.5$
H_0: $\mu_1 = \mu_2$ H_1: $\mu_1 \neq \mu_2$

$$U = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1 = 11(13) + \frac{11(12)}{2} - 134.5 = 74.5$$

$$\mu_U = \frac{n_1 n_2}{2} = \frac{11(13)}{2} = 71.5$$

$$\sigma_U = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}} = \sqrt{\frac{11(13)(25)}{12}} = 17.26$$

Los valores críticos de z son ± 1.645. El valor estandarizado U es

$$z = \frac{U - \mu_U}{\sigma_U} = \frac{74.5 - 71.5}{17.26} = 0.174$$

Debido a que el valor estandarizado U está dentro de los valores críticos, se acepta H_0. No hay una diferencia significativa entre las escuelas.

14.4 Prueba de corridas de una sola muestra

Concepto de aleatoriedad

Hasta ahora, hemos supuesto que las muestras en los problemas se seleccionaron aleatoriamente, esto es, se eligieron sin preferencia o sesgo. ¿Qué pasaría si llegara a observar patrones recurrentes en una muestra elegida por otra persona? Suponga que los solicitantes de capacitación para trabajos especializados tuvieran que seleccionarse sin importar el sexo, a partir de una población grande. Usando la notación M = mujer y H = hombre, encuentra que el primer grupo entra en este orden:

M, M, M, M, H, H, H, H, M, M, M, M, H, H, H, H

Por inspección, concluiría que aunque el número total de solicitantes se divide por igual entre los sexos, el orden no es aleatorio. Un proceso aleatorio rara vez enumeraría dos elementos en grupos alternos de cuatro. Suponga ahora que los solicitantes empiezan a llegar en este orden:

M, H, M, H, M, H, M, H, M, H, M, H, M, H, M, H

Resulta igual de irracional pensar que un proceso de selección aleatorio produciría un patrón de orden como éste de hombres y mujeres. En este caso, también, la *proporción* de mujeres respecto a hombres es correcta, pero sospecharía respecto al *orden* en el que llegaron.

La teoría de corridas

Para permitirnos probar la aleatoriedad del orden de las muestras, los estadísticos han desarrollado la *teoría de corridas*. **Una *corrida* es una secuencia de ocurrencias idénticas precedidas y seguidas de ocurrencias diferentes o del todo por ninguna.** Si los hombres y las mujeres entran de la siguiente manera, la secuencia contendrá tres corridas:

$$\underbrace{\text{M,}}_{\text{1a.}}\ \underbrace{\text{H, H,}}_{\text{2a.}}\ \underbrace{\text{H, H,}}_{\text{3a.}}$$

Y esta secuencia contiene seis corridas:

$$\underbrace{\text{M, M, M,}}_{\text{1a.}}\ \underbrace{\text{H, H,}}_{\text{2a.}}\ \underbrace{\text{M,}}_{\text{3a.}}\ \underbrace{\text{H, H, H, H,}}_{\text{4a.}}\ \underbrace{\text{M, M, M, M,}}_{\text{5a.}}\ \underbrace{\text{H}}_{\text{6a.}}$$

Una *prueba de corridas* usaría los siguientes símbolos si tuviera sólo dos tipos de ocurrencias:

Símbolos empleados para una prueba de corridas

n_1 = número de ocurrencias del tipo 1

n_2 = número de ocurrencias del tipo 2

r = número de corridas

Apliquemos estos símbolos a un patrón diferente para la llegada de los solicitantes:

H, M, M, H, H, H, H, M, M, M, H, H, M, H, M, M, H

En este caso, los valores de n_1, n_2 y r serían

$n_1 = 8$ ← Número de mujeres

$n_2 = 9$ ← Número de hombres

$r = 9$ ← Número de corridas

Un problema que ilustra una prueba de corridas de una sola muestra

Un fabricante de cereal para el desayuno usa una máquina para introducir aleatoriamente uno de dos tipos de muñecos en cada caja. La compañía desea una aleatoriedad tal que no todos los niños de un vecindario terminen con el mismo muñeco. Los probadores eligen muestras de 60 cajas sucesivas para ver si la máquina está mezclando adecuadamente los dos tipos de muñecos. Usando los símbo-

los A y B para representar los dos tipos de muñecos, un probador reportó que uno de estos lotes se presentó como sigue:

B, A, B, B, B, A, A, A, B, B, A, B, B, B, B, A, A, A, A, B,
A, B, A, A, B, B, B, A, A, B, A, A, A, A, B, B, A, B, B, A,
A, A, A, B, B, A, B, B, B, B, A, A, B, B, A, B, A, A, B, B

Los valores de esta prueba serían:

$n_1 = 29$ ← Número de cajas que contienen el muñeco A

$n_2 = 31$ ← Número de cajas que contienen el muñeco B

$r = 29$ ← Número de corridas

La distribución de muestreo del estadístico *r*

El estadístico *r*, base de la prueba de corridas de una sola muestra

El *número de corridas*, *r*, es un estadístico con su propia distribución de muestreo especial y su propia prueba. Es obvio que las corridas pueden ser de diferente longitud y en una muestra pueden ocurrir diversos números de corridas. Los estadísticos pueden probar que demasiadas corridas o muy pocas en una muestra indican que intervino algo más que el azar cuando se seleccionaron los elementos. **Una *prueba de corridas de una sola muestra,* entonces, está basada en la idea de que *muy pocas* o *demasiadas* corridas muestran que los elementos no fueron elegidos aleatoriamente.**

Para derivar la media de la distribución de muestreo del estadístico *r* utilice la siguiente fórmula:

Media y error estándar del estadístico *r*

Media de la distribución muestral del estadístico *r*

$$\mu_r = \frac{2n_1n_2}{n_1 + n_2} + 1 \qquad [14\text{-}6]$$

Al aplicar esta fórmula a la compañía de cereales, la media de la estadística *r* sería:

$$\mu_r = \frac{(2)(29)(31)}{29 + 31} + 1$$

$$= \frac{1{,}798}{60} + 1$$

$$= 29.97 + 1$$

$$= 30.97 \leftarrow \text{Media del estadístico } r$$

El error estándar del estadístico *r* puede calcularse con esta fórmula de aspecto formidable:

Error estándar del estadístico *r*

$$\sigma_r = \sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}} \qquad [14\text{-}7]$$

Para nuestro problema, el error estándar del estadístico *r* se convierte:

$$\sigma_r = \sqrt{\frac{(2)(29)(31)(2 \times 29 \times 31 - 29 - 31)}{(29 + 31)^2(29 + 31 - 1)}}$$

$$= \sqrt{\frac{(1{,}798)(1{,}738)}{(60)^2(59)}}$$

$$= \sqrt{14.71}$$

$$= 3.84 \leftarrow \text{Error estándar del estadístico } r$$

Prueba de las hipótesis

En la prueba de corridas de una sola muestra, la distribución muestral de *r* puede aproximarse de cerca por la distribución normal si n_1 *o bien* n_2 es mayor que 20. Nuestra compañía de cereales tiene una muestra de 60 cajas, así que podemos usar la aproximación normal. La gerencia está interesada en probar, al nivel de 0.20, la hipótesis de que los muñecos están mezclados aleatoriamente, por lo que la prueba se convierte en:

Formulación de las hipótesis

H_0: En una prueba de corridas de una sola muestra, no es adecuado formular las hipótesis con símbolos ← Hipótesis nula: los muñecos están mezclados aleatoriamente

H_1: En una prueba de corridas de una sola muestra, no es adecuado formular las hipótesis con símbolos ← Hipótesis alternativa: los muñecos no están mezclados aleatoriamente

$\alpha = 0.20$ ← Nivel de significancia para probar estas hipótesis

Ilustración de la prueba

Puesto que demasiadas o muy pocas corridas indicarían que el proceso mediante el cual se introducen los muñecos en las cajas no es aleatorio, una prueba de dos colas es apropiada. La figura 14-6 ilustra esta prueba.

A continuación se usa la ecuación 6-2 para *estandarizar* el estadístico de la muestra *r*, 29, restando μ_r, su media y dividiendo entre σ_r, su error estándar.

$$z = \frac{r - \mu_r}{\sigma_r} \qquad [6\text{-}2]$$

$$= \frac{29 - 30.97}{3.84}$$

$$= -0.513$$

FIGURA 14-6

Prueba de hipótesis de dos colas al nivel de significancia de 0.20

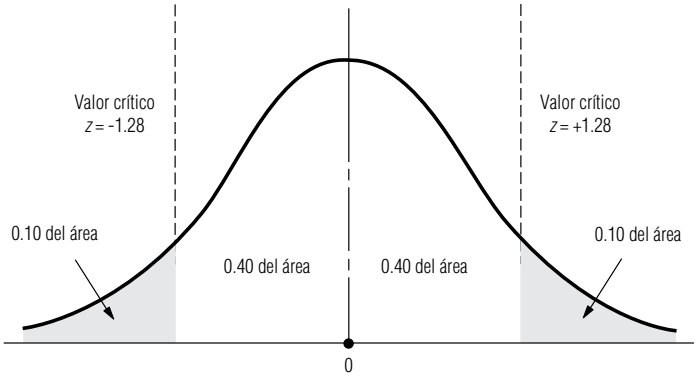
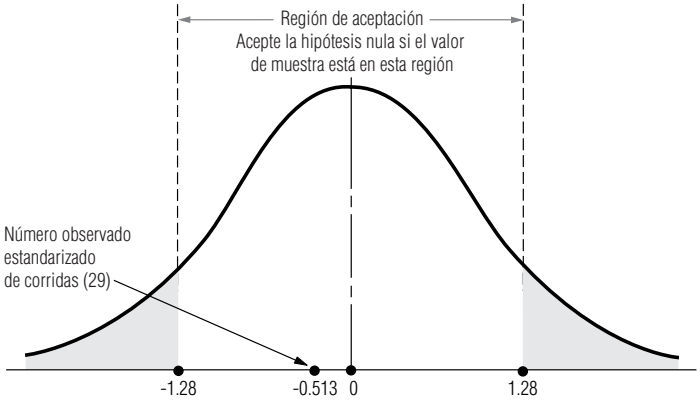


FIGURA 14-7

Prueba de hipótesis de dos colas al nivel de significancia de 0.20, que ilustra la región de aceptación y el número observado de corridas



Colocando el valor estandarizado en la escala *z* de la figura 14-7 se ve que cae dentro de los valores críticos para esta prueba. Por tanto, la administración debe aceptar la hipótesis nula y concluir, a partir de esta prueba, que los muñecos se ponen en las cajas en orden aleatorio.

SUGERENCIAS Y SUPOSICIONES

Las pruebas de corridas se pueden usar de manera efectiva en situaciones de control de calidad. Recordará del capítulo 10 que la variación en la calidad puede ser sistemática o aleatoria, y si una variación es sistemática, se puede corregir. Una prueba de corridas puede detectar los tipos de patrones en la calidad de la producción que están asociados con la variación sistemática. Sugerencia: casi todas las pruebas de corridas son de dos colas porque la pregunta que debe responderse es si hay muy pocas corridas. Recuerde también que las pruebas de corridas usan el estadístico *r* cuya distribución se puede describir bien mediante una distribución normal, *siempre que* n_1 *o* n_2 sea mayor que 20.

Ejercicios 14.4

Ejercicios de autoevaluación

EA 14-5 El profesor Ike Newton desea determinar si sus alumnos más brillantes (aquellos con las mejores notas) tienden a devolver más pronto sus exámenes (porque pueden recordar el material más rápido) o más tarde (porque toman más tiempo para escribir lo que saben) que los demás de la clase. Para cierto examen particular de física, observa que los estudiantes obtienen las siguientes calificaciones de acuerdo con el orden de entrega de sus exámenes.

Orden	Calificaciones									
1-10	94	70	85	89	92	98	63	88	74	85
11-20	69	90	57	86	79	72	80	93	66	74
21-30	50	55	47	59	68	63	89	51	90	88

a) Si el profesor Newton considera a quienes obtienen notas de 90 o más como sus estudiantes brillantes, entonces, a un nivel de significancia del 5%, ¿podría concluir que los estudiantes más brillantes entregaron sus exámenes de manera aleatoria?

b) Si una calificación de 60 o más permite aprobar la materia del profesor Newton, ¿entonces los estudiantes que aprobaron contra los que no lo hicieron entregaron sus exámenes aleatoriamente? (También use un nivel de significancia del 5%.)

Conceptos básicos

■ **14-24** Pruebe la aleatoriedad de la siguiente muestra usando un nivel de significancia de 0.05:

A, B, A, A, A, B, B, A, B, B, A, A, B, A, B, A, A, B, B, B, B, A, B, B,
A, A, A, B, A, B, A, A, B, B, A, B, B, A, A, A, B, B, A, A, B, A, A, A

Aplicaciones

■ **14-25** Se inspeccionó una secuencia de pequeñas esculturas de vidrio para detectar daños ocasionados en el envío. La secuencia de piezas aceptables y dañadas fue la siguiente:

D, A, A, A, D, D, D, D, D, A, A, D, D, A, A, A, A, D, A, A, D, D, D, D

Pruebe la aleatoriedad de los daños ocasionados en el envío usando el nivel de significancia de 0.05.

■ **14-26** El *News and Clarion* mantenía un registro del sexo de las personas que llamaban a la oficina de circulación para quejarse de los problemas con la entrega del periódico dominical. Para un domingo reciente, estos datos fueron los siguientes:

M, F, F, F, M, M, F, M, F, F, F, F, F, M, M, M, F, M, F, M, F, F, F, F, M, M, M, M, M

Usando el nivel de significancia de 0.05, pruebe la aleatoriedad de esta secuencia. ¿Hay algo respecto a la naturaleza de este problema que lo induzca a creer que una secuencia así no es aleatoria?

■ **14-27** La Agencia de Servicios Sociales del condado de Kerwin mantiene este registro del número diario de solicitantes de asesoría matrimonial en el orden en que aparecen en la agencia durante 30 días hábiles.

3, 4, 6, 8, 4, 6, 7, 2, 5, 7, 4, 8, 4, 7, 9, 5, 9, 10,
5, 7, 4, 9, 8, 9, 11, 6, 7, 5, 9, 12

Pruebe la aleatoriedad de esta secuencia viendo si los valores mayores y menores que la media ocurren en orden aleatorio. Use el nivel de significancia de 0.10. ¿Puede pensar en alguna característica del entorno de este problema que apoye lo que encontró con la estadística?

■ **14-28** El dueño de un restaurante ha observado con los años que las parejas de más edad parecen comer más temprano que las parejas jóvenes en su tranquilo y romántico restaurante. Sospecha que tal vez sea porque deben dejar a los niños con alguien que los cuide y porque las parejas de más edad se acuestan más temprano. Una noche, decidió llevar un registro de la llegada de parejas al restaurante. Anotó si cada pareja tenía menos o más que 30 años. Sus notas se reproducen a continuación. (A = 30 y mayores; B = menores de 30.)

(5:30 p.m.) A, A, A, A, A, A, B, A, A, A, A, A, A, B, B,
B, A, B, B, B, B, B, B, A, B, B, B, A, B, B, B (10 p.m.)

Para un nivel de significancia del 5%, ¿estaba en lo correcto el dueño del restaurante en cuanto a que la edad de sus clientes a distintas horas de la noche no es aleatoria?

■ **14-29** Kathy Phillips está a cargo de la programación de producción en una imprenta. La compañía tiene seis prensas grandes, que a menudo se descomponen, y uno de los mayores problemas de Kathy es cumplir con las fechas de entrega cuando ocurren desperfectos imprevistos en las prensas. Sospecha que cada la semana las viejas prensas se descomponen antes que las nuevas, ya que todas las prensas se verifican y reparan durante el fin de semana. Para probar su hipótesis, Kathy registró el número de las prensas que fallaron durante la semana. Las prensas 1, 2 y 3 son las más antiguas.

Número de prensa en el orden de falla

1, 2, 3, 1, 4, 5, 3, 1, 2, 5, 1, 3, 6, 2, 3, 6, 2, 2, 3, 5, 4,
6, 4, 2, 1, 3, 4, 5, 5, 1, 4, 5, 2, 3, 5, 6, 4, 3, 2, 5, 4, 3

a) A un nivel de significancia del 5%, ¿tiene Kathy una hipótesis válida respecto a que las fallas de las prensas no son aleatorias?
b) ¿Es apropiada su hipótesis en cuanto a la decisión que desea tomar de reprogramar más trabajo a principios de la semana en las prensas más nuevas?

■ **14-30** Martha Bowen, gerente de departamento de una gran compañía de investigación de mercados, está a cargo de todos los análisis de datos de investigación de la compañía. La exactitud y minuciosidad son su responsabilidad. El departamento emplea varios ayudantes de investigación para que hagan algunos análisis y usa una computadora para otros análisis. Por lo general, cada semana Martha elige aleatoriamente los análisis terminados antes de entregarlos y hace pruebas para asegurar que son correctos y exhaustivos. La ayudante de Martha, Kim Tadlock elige al azar 49 análisis por semana entre los terminados y archivados diariamente, y Martha hace los análisis de nuevo. Ella desea cerciorarse de que el proceso de selección es aleatorio, con el fin de poder asegurar que los análisis de computadora y los hechos a mano se verifican periódicamente. Para ello, dispuso que los ayudantes de investigación colocaran una marca especial en la parte de atrás de los registros, de manera que pudieran identificarse. Kim, no estaba al tanto de la marca, para que la aleatoriedad de la prueba no se viera afectada. Kim completó su muestra con los siguientes datos:

Muestras de análisis de datos para una semana
(1, por computadora; 2, a mano)

1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 2, 1, 1, 1, 1, 1,
1, 1, 1, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1

a) A un nivel de significancia del 1%, ¿podría concluir que la muestra fue aleatoria?
b) Si la muestra estuviera distribuida como sigue, ¿sería aleatoria?

1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2

c) Puesto que los análisis por computadora son mucho más rápidos que los hechos a mano, y como es posible hacer varios de los análisis por computadora, cada semana hay aproximadamente tres veces más análisis de éstos que los hechos a mano. ¿Existe evidencia estadística en el inciso a) para apoyar la creencia de que en alguna parte del proceso de muestreo no ocurre la aleatoriedad? Si es así, ¿cuál es la evidencia?

d) ¿La conclusión a la que llegó en el inciso c) lo lleva a alguna nueva conclusión respecto a la prueba de corridas de una sola muestra, particularmente en lo que se refiere a su respuesta en a)?

■ **14-31** Bank of America tiene curiosidad acerca del grado académico de los usuarios de sus cajeros automáticos (ATM) colocados en el edificios de la Sociedad de Alumnos. Los estudiantes de primero y segundo año se clasifican como tipo A, los de tercero y cuarto como tipo B. Los datos de 45 personas que usaron el ATM el viernes en la tarde son los siguientes. Pruebe que esta secuencia sea aleatoria para un nivel de significancia de 0.05.

B B B A A A B A A A A A A B B B B A B A A A A B B A A B B B B A B B B B A A A A A A B B B

■ **14-32** El First National Bank de Smithville registró el sexo de los primeros 40 clientes, que aparecieron el pasado martes, con esta notación:

M, F, M, M, M, M, F, F, M, M, M, F, M, M, M, M, M, F, F, M,
F, M, M, M, F, M, M, M, M, M, M, F, M, M, M, M, M, F, F, M

Al nivel de significancia de 0.05, pruebe la aleatoriedad de esta secuencia. ¿existe algo en el área bancaria o en la naturaleza de este problema que lo lleve a aceptar intuitivamente lo que encontró estadísticamente?

Soluciones a los ejercicios de autoevaluación

EA 14-5 a) G denota a los estudiantes con 90 o más y L denota a los que tienen menos de 90:

$$G\,L\,L\,L\,G\,G\,L\,L\,L\,L\,L\,G\,L\,L\,L\,L\,L\,G\,L\,L\,L\,L\,L\,L\,L\,L\,L\,G\,L$$

$n_1 = \#$ de G´$s = 6$ $\qquad r = 10$

$n_2 = \#$ de L´$s = 24$ $\qquad \alpha = 0.05$

$$\mu_r = \frac{2n_1n_2}{n_1 + n_2} + 1 = \frac{2(6)(24)}{30} + 1 = 10.6$$

$$\sigma_r = \sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}} = \sqrt{\frac{2(6)(24)[2(6)(24) - 6 - 24]}{(30)^2(29)}}$$
$$= 1.69$$

Los valores críticos de z son ±1.96. El valor estandarizado de r es

$$z = \frac{r - \mu_r}{\sigma_r} = \frac{10 - 10.6}{1.69} = -0.355$$

de manera que se acepta H_0. La secuencia es aleatoria.

b) Si P denota aprobar (≥ 60) y F denota no aprobar (< 60), se obtiene

$$P\,P\,P\,P\,P\,P\,P\,P\,P\,P\,P\,P\,P\,F\,P\,P\,P\,P\,P\,P\,P\,P\,P\,F\,F\,F\,F\,P\,P\,P\,F\,P\,P$$

$n_1 = \#$ de P´$s = 24$ $\qquad r = 7$

$n_2 = \#$ de F´$s = 6$ $\qquad \alpha = 0.05$

$$\mu_r = \frac{2(24)(6)}{30} + 1 = 10.6$$

$$\sigma_r = \sqrt{\frac{2(24)(6)[2(24)(6) - 24 - 6]}{(30)^2(29)}} = 1.69$$

Los valores críticos de z son ±1.96. El valor estandarizado de r es

$$z = \frac{7 - 10.6}{1.69} = -2.13$$

de manera que se rechaza H_0 porque $z < -1.96$. Esta secuencia no es aleatoria.

14.5 Correlación de rango

Función del coeficiente de correlación de rango

Los capítulos 12 y 13 nos introdujeron a los conceptos de correlación y coeficiente de correlación, una medida de la cercanía de asociación entre dos variables. Con frecuencia, en el análisis de correlación, la información no está disponible en forma de valores numéricos como los utilizados en los problemas de esos capítulos. Sin embargo, si podemos asignar clasificaciones a los elementos de cada una de las dos variables que estamos estudiando, entonces podemos calcular un *coeficiente de correlación de rango*. **Éste es una medida de la correlación que existe entre los dos conjuntos de rangos, una medida del grado de asociación entre las variables que no podríamos calcular de otra manera.**

Otra ventaja del uso de la correlación de rango

Una segunda razón para aprender el método de correlación de rango es la posibilidad de simplificar el proceso de cálculo de un coeficiente de correlación a partir de un conjunto de datos muy grande para cada una de las dos variables. Para probar lo tedioso que esto puede ser, intente aumentar uno de los problemas de correlación del capítulo 12, por un factor de 10, y realizar los cálculos necesarios. En lugar de tener que hacer estos cálculos, podemos calcular una medida de asociación basada en los *rangos* de las observaciones, *no en los valores numéricos* de los datos. Esta medición se conoce como coeficiente de correlación de rango de Spearman, en honor al estadístico que lo desarrolló a principios del siglo XX.

El coeficiente de correlación de rangos

Listado de las variables clasificadas

Mediante un par de ejemplos, aprenderemos a calcular e interpretar esta medida de la asociación entre dos variables clasificadas. Primero, consideremos la tabla 14-9, que enumera cinco personas y compara el rango o clasificación académica que obtuvieron en la universidad con el nivel que han logrado en cierta compañía 10 años después de graduarse. El valor 5 representa el rango más alto del grupo; 1 es el más bajo.

Cálculo del coeficiente de correlación de rangos

Usando la información de la tabla 14-9, podemos calcular un coeficiente de correlación de rangos entre el éxito en la universidad y el nivel logrado en la compañía 10 años después. Lo único que necesitamos es la ecuación 14-8 y unos cuantos cálculos.

Coeficiente de correlación de rangos

$$r_s = 1 - \frac{6\Sigma d^2}{n(n_2 - 1)} \qquad [14\text{-}8]$$

donde,

- r_s = coeficiente de correlación de rangos (observe que el subíndice s, de Spearman, distingue esta r de la calculada en el capítulo 12)
- n = número de observaciones apareadas
- Σ = notación que significa “la suma de”
- d = diferencia entre los rangos para cada pareja de observaciones

Tabla 14-9
Comparación de los rangos de cinco estudiantes

Estudiante	Rango universitario	Rango en la compañía 10 años después
John	4	4
Margaret	3	3
Debbie	1	1
Steve	2	2
Lisa	5	5

Los cálculos en forma tabular son sencillos, como se ve en la tabla 14-10. Por tanto, tenemos toda la información que necesitamos para encontrar el coeficiente de correlación de rango para este problema:

$$r_s = 1 - \frac{6\Sigma d^2}{n(n^2 - 1)} \qquad [14\text{-}8]$$

$$= 1 - \frac{6(0)}{5(25 - 1)}$$

$$= 1 - \frac{0}{120}$$

$$= 1 \leftarrow \text{Coeficiente de correlación de rangos}$$

Explicación de los valores del coeficiente de correlación de rango

Como estudiamos en el capítulo 12, este coeficiente de correlación de 1 indica que existe una asociación perfecta o *correlación perfecta* entre las dos variables. Esto verifica lo que se ve en la tabla 14-9, que los rangos universitarios y de la compañía para cada persona son idénticos.

Cálculo de otro coeficiente de correlación de rangos

Otro ejemplo nos familiarizará más con el manejo del coeficiente de correlación de rangos. En la tabla 14-11 se presenta el caso de otras cinco personas, pero esta vez los rangos universitarios y de la compañía 10 años después parecen ser opuestos. Podemos calcular la diferencia entre los rangos

Tabla 14-10
Generación de información para calcular el coeficiente de correlación de rangos

Estudiante	Rango universitario (1)	Rango en la compañía (2)	Diferencia entre los dos rangos (1) − (2)	Diferencia al cuadrado $[(1) - (2)]^2$
John	4	4	0	0
Margaret	3	3	0	0
Debbie	1	1	0	0
Steve	2	2	0	0
Lisa	5	5	0	0
				$\Sigma d^2 = \mathbf{0}$ ← Suma de los cuadrados de las diferencias

Tabla 14-11
Generación de datos para calcular el coeficiente de correlación de rangos

Estudiante	Rango universitario (1)	Rango en la compañía (2)	Diferencia entre los dos rangos (1) − (2)	Diferencia al cuadrado $[(1) - (2)]^2$
JohRoy	5	1	4	16
David	1	5	−4	16
Jay	3	3	0	0
Charlotte	2	4	−2	4
Kathy	4	2	2	4
				$\Sigma d^2 = \mathbf{40}$ ← Suma de los cuadrados de las diferencias

de cada pareja de observaciones, encontrar d^2, y luego sumar todas las d^2. Sustituyendo estos valores en la ecuación 14-8, encontramos un coeficiente de correlación de rango cuyo valor es −l:

$$r_s = 1 - \frac{6\Sigma d^2}{n(n^2 - 1)} \qquad [14\text{-}8]$$

$$= 1 - \frac{6(40)}{5(25 - 1)}$$

$$= 1 - \frac{240}{120}$$

$$= 1 - 2$$

$$= -1 \leftarrow \text{Coeficiente de correlación de rangos}$$

Interpretación de los resultados

En el capítulo 12, aprendimos que un coeficiente de correlación de −1 representa una *correlación inversa perfecta*. Eso es justamente lo que ocurre en nuestro caso: la gente que más destacó en la universidad terminó en los rangos menores de una organización. Ahora apliquemos estas ideas.

Solución de un problema usando correlación de rangos

La correlación de rangos es una técnica útil para examinar a conexión entre calidad del aire y la evidencia de enfermedades pulmonares que analizamos en el problema del inicio del capítulo. La tabla 14-12 reproduce los datos encontrados por la organización de la salud que estudia el problema. En la misma tabla, hacemos algunos cálculos necesarios para encontrar r_s.

Búsqueda del coeficiente de correlación de rangos

Usando los datos de la tabla 14-12 y la ecuación 14-8, podemos encontrar el coeficiente de correlación de rangos para este problema:

$$r_s = 1 - \frac{6\Sigma d^2}{n(n^2 - 1)} \qquad [14\text{-}8]$$

$$= 1 - \frac{6(58)}{11(121 - 1)}$$

$$= 1 - \frac{348}{1{,}320}$$

$$= 1 - 0.2636$$

$$= 0.7364 \leftarrow \text{Coeficiente de correlación de rangos}$$

Interpretación de los resultados

Un coeficiente de correlación de 0.736 sugiere una asociación positiva sustancial entre calidad promedio del aire y la ocurrencia de enfermedades pulmonares, al menos en las 11 ciudades muestreadas; esto es, altos niveles de contaminación van acompañados de enfermedades pulmonares.

¿Cómo podemos probar este valor de 0.736? Podemos aplicar los mismos métodos usados para probar hipótesis en los capítulos 8 y 9. Al realizar esas pruebas en r_s, tratamos de evitar el error de concluir que existe una asociación entre dos variables si en realidad no existe en la población de la cual se tomaron esas dos muestras, es decir, si el coeficiente de correlación de rango de la *población*, ρ_s (*ro sub s*), es realmente igual a cero.

Prueba de hipótesis acerca de la correlación de rangos

Para pequeños valores de *n* (*n* menor o igual que 30), la distribución de r_s no es normal, y a diferencia de otras estadísticas de muestras pequeñas que hemos encontrado, no es apropiado usar la distribución *t* para probar hipótesis acerca del coeficiente de correlación de rangos. En su lugar, utilizamos la tabla 7 del apéndice para determinar las regiones de aceptación y de rechazo de las hipótesis. En nuestro problema actual, suponga que la organización de salud desea probar, al

Tabla 14-12
Rangos de 11 ciudades

Ciudad	Rango de la calidad de aire (1)	Rango de enfermedad pulmonar (2)	Diferencia entre los dos rangos (1) − (2)	Diferencia al cuadrado $[(1) - (2)]^2$
A	4	5	−1	1
B	7	4	3	9
C	9	7	2	4
D	1	3	−2	4
E	2	1	1	1
F	10	11	−1	1
G	3	2	1	1
H	5	10	−5	25
I	6	8	−2	4
J	8	6	2	4
K	11	9	2	4
	Mejor rango = 11 Peor rango = 1			$\Sigma d^2 = \overline{58}$ ← Suma de los cuadrados de las diferencias

nivel de significancia de 0.05, la hipótesis nula de que existe una correlación de cero en los datos clasificados de *todas* las ciudades del mundo. Nuestro problema entonces se convierte en:

Formulación de las hipótesis

$H_0: \rho_s = 0$ ← Hipótesis nula: no existe correlación en los datos clasificados de la población

$H_1: \rho_s \neq 0$ ← Hipótesis alternativa: existe correlación en los datos clasificados de las poblaciones

$\alpha = 0.05$ ← Nivel de significancia para probar estas hipótesis

Una prueba de dos colas es adecuada, así que consultamos la tabla 7 del apéndice, en el renglón correspondiente a $n = 11$ (el número de ciudades) y la columna correspondiente al nivel de significancia de 0.05. Encontramos que los valores críticos para r_s son ± 0.6091, esto es, el límite superior de la región de aceptación es 0.6091 y el límite inferior es -0.6091.

La figura 14-8 ilustra los límites de la región de aceptación y el coeficiente de correlación de rangos calculado a partir de la muestra de calidad de aire. De esta figura, podemos ver que el coeficiente de correlación de rango cae fuera de la región de aceptación. Por consiguiente, rechazamos la hipótesis nula de no correlación y concluimos que existe una asociación entre los niveles de calidad del aire y la incidencia de enfermedades pulmonares en las ciudades del mundo.

La distribución apropiada para valores de *n* mayores que 30

Si el tamaño de la muestra es mayor que 30, ya no podemos utilizar la tabla 7 del apéndice. Sin embargo, cuando *n* es mayor que 30, la distribución muestral de r_s es aproximadamente normal, con media de cero y desviación estándar de $1/\sqrt{n-1}$. En consecuencia, el error estándar de r_s es

Error estándar del coeficiente de la correlación de rangos

$$\sigma_{r_s} = \frac{1}{\sqrt{n-1}} \qquad [14\text{-}9]$$

y podemos usar la tabla 1 del apéndice para encontrar los valores de *z* para probar las hipótesis con respecto a la correlación de rangos de población.

Ejemplo con *n* mayor que 30

Como un ejemplo de prueba de hipótesis de coeficientes de correlación de rango cuando *n* es mayor que 30; consideremos el caso de un científico social que intenta determinar si la gente brillante tiende a elegir un cónyuge que también sea brillante. Selecciona aleatoriamente 32 parejas y prueba para ver si existe una correlación de rangos significativa en los coeficientes intelectuales (CI) de las parejas. Sus datos y cálculos se dan en la tabla 14-13.

FIGURA 14-8

Prueba de hipótesis de dos colas, usando la tabla 7 del apéndice al nivel de significancia de 0.05, donde se ve la región de aceptación y el coeficiente de correlación de rangos muestrales

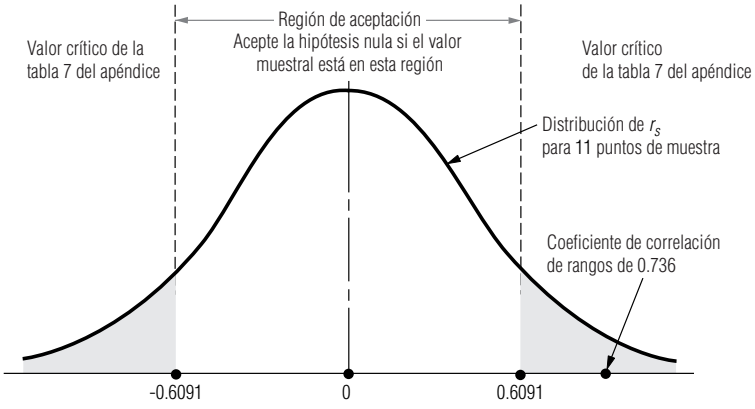


Tabla 14-13

Cálculo de la correlación de rangos del CI de esposos y esposas

Pareja (1)	CI del esposo (2)	CI de la esposa (3)	Rango del esposo (4)	Rango de la esposa (5)	Diferencia entre rangos (4) − (5)	Diferencia al cuadrado $[(4) - (5)]^2$
1	95	95	8	4.5	3.5	12.25
2	103	98	20	8.5	11.5	132.25
3	111	110	26	23	3	9.00
4	92	88	4	2	2	4.00
5	150	106	32	18	14	196.00
6	107	109	24	21.5	2.5	6.25
7	90	96	3	6	−3	9.00
8	108	131	25	32	−7	49.00
9	100	112	17.5	25.5	−8	64.00
10	93	95	5.5	4.5	1	1.00
11	119	112	29	25.5	3.5	12.25
12	115	117	28	30	−2	4.00
13	87	94	1	3	−2	4.00
14	105	109	21	21.5	−0.5	0.25
15	135	114	31	27	4	16.00
16	89	83	2	1	1	1.00
17	99	105	14.5	16.5	−2	4.00
18	106	115	22.5	28	−5.5	30.25
19	126	116	30	29	1	1.00
20	100	107	17.5	19	−1.5	2.25
21	93	111	5.5	24	−18.5	342.25
22	94	98	7	8.5	−1.5	2.25
23	100	105	17.5	16.5	1	1.00
24	96	103	10	15	−5	25.00
25	99	101	14.5	13	1.5	2.25
26	112	123	27	31	−4	16.00
27	106	108	22.5	20	2.5	6.25
28	98	97	12.5	7	5.5	30.25
29	96	100	10	11.5	−1.5	2.25
30	98	99	12.5	10	2.5	6.25
31	100	100	17.5	11.5	6	36.00
32	96	102	10	14	−4	16.00
					Suma de cuadrados de las diferencias → $\Sigma d^2 =$	**1,043.50**

Usando los datos de la tabla 14-13 y la ecuación 14-8, podemos encontrar el coeficiente de correlación de rangos para este problema:

$$r_s = 1 - \frac{6\Sigma d^2}{n(n^2 - 1)} \quad [14\text{-}8]$$

$$= 1 - \frac{6(1{,}043.5)}{32(1{,}024 - 1)}$$

$$= 1 - \frac{6{,}261}{32{,}736}$$

$$= 1 - 0.1913$$

$$= 0.8087 \leftarrow \text{Coeficiente de correlación de rangos}$$

Si el científico social desea probar su hipótesis al nivel de significancia de 0.01, su problema puede formularse así:

Formulación de las hipótesis

H_0: $\rho_s = 0$ ← Hipótesis nula: no existe correlación de rangos en la población, esto es, la inteligencia de los esposos y esposas está mezclada al azar

H_1: $\rho_s > 0$ ← Hipótesis alternativa: la correlación de rangos de la población es positiva, es decir, las personas brillantes eligen esposos brillantes

$\alpha = 0.01$ ← Nivel de significancia para probar estas hipótesis

Es adecuado realizar una prueba de cola superior. De la tabla 1 del apéndice, encontramos que el valor *z* crítico para el nivel de significancia de 0.01 es 2.33. La figura 14-9 ilustra esta prueba de hipótesis, donde la región sombreada en la cola superior de la distribución corresponde al nivel de significancia de 0.01.

Para calcular nuestro estadístico de prueba, primero encontramos el error estándar r_s:

$$\sigma_{r_s} = \frac{1}{\sqrt{n - 1}} \quad [14\text{-}9]$$

$$= \frac{1}{\sqrt{32 - 1}} = 0.1796$$

Ahora podemos usar la ecuación 6-2 para *estandarizar* el coeficiente de correlación del rango, r_s, restando 0, su valor hipotético y dividiendo entre σ_{r_s}, su error estándar.

$$z = \frac{r_s - 0}{\sigma_{r_s}} \quad [6\text{-}2]$$

$$= \frac{0.8087}{0.1796}$$

$$= 4.503$$

Interpretación de los resultados

La figura 14-10 muestra el límite de la región de aceptación y el coeficiente de correlación de rangos que calculamos de los datos del CI. En esta figura, podemos ver que el coeficiente de correlación de rango de 0.809 cae bastante fuera de la región de aceptación. Por tanto, rechazamos la hipótesis nula de no correlación y concluimos que la gente brillante tiende a elegir esposos brillantes.

Una propiedad especial de la correlación de rangos

Ventaja de la correlación de rangos

La correlación de rango tiene una ventaja útil sobre el método de correlación que analizamos en el capítulo 12. Suponga que tenemos casos en los que existen una o varias observaciones sumamente

FIGURA 14-9

Prueba de hipótesis de cola superior al nivel de significancia de 0.01

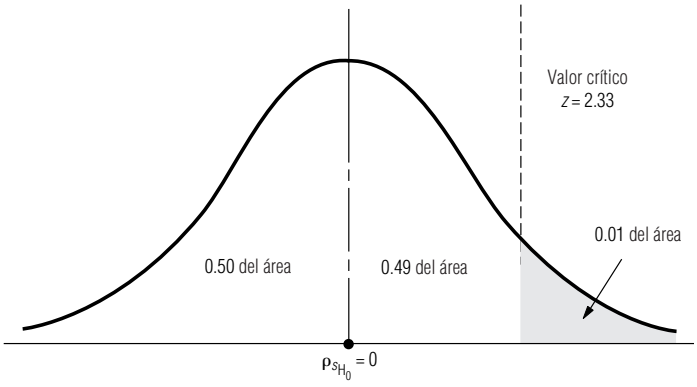
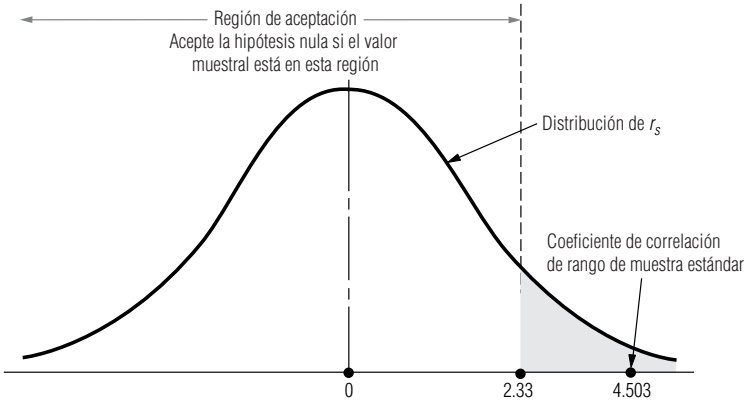


FIGURA 14-10

Prueba de hipótesis de cola superior al nivel de significancia de 0.01; indica la región de aceptación y el coeficiente de correlación de rango de la muestra



extremas en los datos originales. Calcular el coeficiente de correlación con valores numéricos, como lo hicimos en el capítulo 12, puede no producir una buena descripción de la asociación existente entre dos variables. Sin embargo, las observaciones extremas en una prueba de correlación de rangos nunca producirán una diferencia de rangos grande.

Considere el siguiente arreglo de datos de dos variables, *X* y *Y*:

X	10	13	16	19	25
Y	34	40	45	51	117

Debido al valor alto del quinto término de *Y*, obtendríamos dos respuestas significativamente diferentes para *r* si se usan los métodos convencional y de correlación de rangos. En este caso, el método de correlación de rangos es menos sensible al valor extremo. Asignamos un orden de rango de 5 al valor numérico de 117 y evitamos el efecto sobre el valor del coeficiente de correlación.

SUGERENCIAS Y SUPOSICIONES

Cuando existen valores extremos en los datos originales, la correlación de rangos puede producir resultados más útiles que el método de correlación explicado en el capítulo 12, porque las observaciones extremas nunca producen una gran diferencia en el rango. Sugerencia: la correlación de rangos es muy útil cuando los datos no siguen una distribución normal. Tome el caso de una colecta de fondos para una universidad donde se obtienen algunas donaciones "grandes", muchas donaciones menores que $100 y un intervalo muy amplio en medio. No tiene sentido usar las técnicas de correlación del capítulo 12 para investigar las relaciones entre el número de cartas de petición y el tamaño de las donaciones con este tipo de distribución porque las donaciones de millón y medio de dólares distorsionan el estudio. Aplicar la correlación de rangos en un caso como éste funciona bastante bien.

Ejercicios 14.5

Ejercicios de autoevaluación

EA 14-6 Las siguientes son clasificaciones de agresividad (X) y cantidad de ventas en el último año (Y) de ocho vendedores. ¿Existe una correlación significativa entre las dos mediciones? Utilice el nivel de significancia de 0.10.

X	30	17	35	28	42	25	19	29
Y	35	31	43	46	50	32	33	42

Aplicaciones

■ **14-33** Los siguientes son años de experiencia (X) y el promedio de clientes satisfechos (Y) para 10 proveedores de servicios. ¿Existe una correlación de rangos significativa entre las dos medidas? Use un nivel de significancia de 0.05.

X	6.3	5.8	6.1	6.9	3.4	1.8	9.4	4.7	7.2	2.4
Y	5.3	8.6	4.7	4.2	4.9	6.1	5.1	6.3	6.8	5.2

■ **14-34** Un supervisor de planta clasificó una muestra de ocho trabajadores según el número de horas extras trabajadas y la antigüedad en el empleo. ¿La correlación de rangos entre las dos mediciones es significativa al nivel de 0.01?

Cantidad de horas extra	5.0	8.0	2.0	4.0	3.0	7.0	1.0	6.0
Años de empleo	1.0	6.0	4.5	2.0	7.0	8.0	4.5	3.0

■ **14-35** Muchas personas creen que la experiencia gerencial produce mejores relaciones interpersonales entre un gerente y sus empleados. La Quail Corporation tiene los siguientes datos que asocian los años de experiencia del gerente con el número de quejas reportadas el año pasado por los empleados contra ese gerente. Al nivel de significancia de 0.05, ¿la correlación de rangos entre estos dos factores sugiere que la experiencia mejora las relaciones?

Años de experiencia	7	18	17	4	21	27	20	14	15	10
Número de quejas	5	2	4	4	3	2	4	5	4	6

■ **14-36** La Administración de Salud y Seguridad Ocupacional (OSHA, *Occupational Safety and Health Administration*) realizaba un estudio de las relaciones entre los gastos en la seguridad de la planta y la tasa de accidentes en las mismas, en la industria química sintética. Para ajustar la diferencia en tamaño que existía entre algunas de las plantas, OSHA convirtió sus datos en gastos por empleado de producción. Los siguientes son los resultados:

Gastos de compañías químicas por empleado de producción en relación con accidentes por año

Compañía	A	B	C	D	E	F	G	H	I	J	K
Gastos	$60	$37	$30	$20	$24	$42	$39	$54	$48	$58	$26
Accidentes	2	7	6	9	7	4	8	2	4	3	8

¿Existe una correlación significativa entre los gastos y los accidentes en las plantas de las compañías químicas? Use una correlación de rangos (con 1 para el mayor gasto y tasa de accidentes) para respaldar su conclusión. Pruebe al nivel de significancia del 1%.

■ **14-37** Dos profesores de una escuela de administración discutían lo difícil que es predecir el éxito de los graduados sólo basándose en las calificaciones. Un profesor pensaba que el número de años de experiencia que los estudiantes de maestría en administración de empresas tuvieran antes de regresar por su posgrado tal vez era el mejor pronosticador. Usando los siguientes datos, al nivel de significancia de 0.02, ¿qué correlación de rangos es un mejor pronosticador del éxito profesional?

Años de experiencia	4	3	4	3	6	7	1	5	5	2
Promedio de calificaciones	3.4	3.2	3.5	3.0	2.9	3.4	2.5	3.9	3.6	3.0
Rango de éxito (10 = tope)	4	2	6	5	7	9	1	8	10	3

■ **14-38** La Carolina Lighting Company tiene dos entrevistadores de personal capacitados para reclutar gerentes principiantes para encargarse de tiendas nuevas. Aunque cada entrevistador tiene su propio estilo, se considera que ambos son buenos jueces preliminares del potencial gerencial. La gerente de personal se preguntaba qué tanto coincidirían los entrevistadores, así que hizo que ambos evaluaran a 14 solicitantes de manera independiente. Clasificaron a los solicitantes en términos de su grado de contribución potencial a la compañía. Los resultados se presentan a continuación. Use la correlación de rangos y un nivel de significancia del 2% para determinar si existe una correlación positiva significativa entre las dos clasificaciones de los entrevistadores.

Solicitante	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Entrevistador 1	1	11	13	2	12	10	3	4	14	5	6	9	7	8
Entrevistador 2	4	12	11	2	14	10	1	3	13	8	6	7	9	5

■ **14-39** Nancy McKenzie, supervisora de un proceso de ensamble de una cámara litográfica, siente que mientras más tiempo trabaja en equipo un grupo de empleados, mayor es la tasa de producción diaria. Ha recabado los siguientes datos para un grupo de empleados que trabajaron juntos durante 10 días.

Día	1	2	3	4	5	6	7	8	9	10
Producción diaria	4.0	7.0	5.0	6.0	8.0	2.0	3.0	0.5	9.0	6.0

¿Puede concluir Nancy, a un nivel de significancia del 5%, que no existe correlación entre el número de días que trabajaron juntos y la producción diaria?

■ **14-40** Una compañía electrónica que recluta muchos ingenieros, se pregunta si el costo de los esfuerzos extensivos de reclutamiento vale la pena. Si la compañía pudiera confiar (usando un nivel de significancia del 1%) en que la correlación de rangos de población entre el currículum individual de los solicitantes clasificados por el departamento de personal y las calificaciones de las entrevistas es positiva, se sentiría justificado en descontinuar las entrevistas y basarse en las calificaciones por currículum en la contratación. La compañía ha extraído una muestra de 35 solicitantes en los últimos dos años. Sobre la base de la muestra anterior, ¿debe la compañía descontinuar las entrevistas y usar las calificaciones del currículum individual para contratar?

Individuo	Calificación de entrevista	Calificación de currículum	Individuo	Calificación de entrevista	Calificación de currículum
1	81	113	19	81	111
2	88	88	20	84	121
3	55	76	21	82	83
4	83	129	22	90	79
5	78	99	23	63	71
6	93	142	24	78	108
7	65	93	25	73	68
8	87	136	26	79	121
9	95	82	27	72	109
10	76	91	28	95	121
11	60	83	29	81	140
12	85	96	30	87	132
13	93	126	31	93	135
14	66	108	32	85	143
15	90	95	33	91	118
16	69	65	34	94	147
17	87	96	35	94	138
18	68	101			

■ **14-41** Los siguientes son los salarios y edades de los 10 candidatos a doctorado que se gradúan este año de la Escuela de Contabilidad de la Northwest University. Al nivel de significancia de 0.05, ¿la correlación de rangos de edad y salario sugiere que los candidatos de mayor edad obtienen salarios iniciales mayores?

Salario	Edad	Salario	Edad
$67,000	29	55,000	31
60,000	25	59,500	32
57,500	30	63,000	38
59,500	35	69,500	28
50,000	27	72,000	34

■ **14-42** Dee Boone opera un taller de reparación de motores de aeronaves ligeras. Está interesado en mejorar sus estimaciones del tiempo de reparación requerido y cree que el mejor pronosticador es el número de horas de operación del motor desde su última reparación importante. La tabla contiene los datos de 10 motores que Dee reparó recientemente. Al nivel de significancia de 0.10, ¿la correlación de rango sugiere una relación fuerte?

Motor	Horas desde la última reparación importante	Horas requeridas para la reparación
1	1,000	40
2	1,200	54
3	900	41
4	1,450	60
5	2,000	65
6	1,300	50
7	1,650	42
8	1,700	65
9	500	43
10	2,100	66

Soluciones a los ejercicios de autoevaluación

EA 14-6

***X*(rangos)**	6	1	7	4	8	3	2	5
***Y*(rangos)**	4	1	6	7	8	2	3	5
d	2	0	1	−3	0	1	−1	0
d^2	4	0	1	9	0	1	1	0

$$\Sigma d^2 = 16 \qquad n = 8 \qquad \alpha = 0.10$$

$$H_0: \rho_s = 0 \qquad H_1: \rho_s \neq 0$$

$$r_s = 1 - \frac{6\Sigma d^2}{n(n^2 - 1)} = 1 - \frac{6(16)}{8(63)} = 0.8095$$

De la tabla 7 del apéndice, los valores críticos para r_s son ± 0.6190. Como $0.8095 > 0.6190$, se rechaza H_0. La correlación es significativa.

14.6 Prueba de Kolmogorov-Smirnov

La prueba de K-S y sus ventajas

La prueba de Kolmogorov-Smimov, bautizada así en honor de los estadísticos A. N. Kolmogorov y N. V. Smirnov que la desarrollaron, se trata de un método no paramétrico sencillo para probar si existe una diferencia significativa entre una distribución de frecuencias observada y un distribución de frecuencias teórica. La prueba de K-S es, por consiguiente, otra medida de la *bondad de ajuste* de una distribución de frecuencia teórica, como lo es la prueba ji-cuadrada que estudiamos en el capítulo 11. Sin embargo, la prueba de K-S tiene varias ventajas sobre la prueba χ^2: es una prueba más poderosa, y es más fácil de usar, puesto que no requiere que los datos se agrupen de alguna manera.

Una ventaja especial

El estadístico de K-S, D_n, es particularmente útil para juzgar qué tan cerca está la distribución de frecuencias observada de la distribución de frecuencias esperada, porque la distribución de probabilidad de D_n depende del tamaño de muestra n, pero es independiente de la distribución de frecuencias esperada (D_n es un estadístico de "distribución libre").

Un problema que ilustra la prueba de K-S

Suponga que la central telefónica del condado de Orange, en California, ha llevado un registro del número de "transmisores" (un tipo de equipo automático que se utiliza en las centrales telefónicas) usados en un instante dado. Las observaciones se hicieron en 3,754 ocasiones distintas. Para propó-

Tabla 14-14
Frecuencias acumuladas observadas y relativas

Número ocupado	Frecuencia observada	Frecuencia acumulada observada	Frecuencia acumulada observada relativa
0	0	0	0.0000
1	5	5	0.0013
2	14	19	0.0051
3	24	43	0.0115
4	57	100	0.0266
5	111	211	0.0562
6	197	408	0.1087
7	278	686	0.1827
8	378	1,064	0.2834
9	418	1,482	0.3948
10	461	1,943	0.5176
11	433	2,376	0.6329
12	413	2,789	0.7429
13	358	3,147	0.8383
14	219	3,366	0.8966
15	145	3,511	0.9353
16	109	3,620	0.9643
17	57	3,677	0.9795
18	43	3,720	0.9909
19	16	3,736	0.9952
20	7	3,743	0.9971
21	8	3,751	0.9992
22	3	3,754	1.0000

sitos de planeación de Inversión de capital, el funcionario de presupuesto de esta compañía piensa que el patrón de uso sigue una distribución de Poisson con una media de 8.5. Si desea probar esta hipótesis al nivel de significancia de 0.01, puede emplear la prueba de K-S.

Formularíamos nuestras hipótesis de la siguiente manera:

Formulación de las hipótesis

H_0: una distribución de Poisson con $\lambda = 8.5$ es una buena descripción del patrón de uso ← Hipótesis nula

H_1: una distribución de Poisson con $\lambda = 8.5$ no es una buena descripción del patrón de uso ← Hipótesis alternativa

$\alpha = 0.01$ ← Nivel de significancia para probar estas hipótesis

Cálculo y comparación de frecuencias esperadas

A continuación, enumeraremos los datos observados. La tabla 14-14 lista las frecuencias observadas y las transforma en frecuencias acumuladas observadas relativas.

Ahora podemos utilizar la fórmula de Poisson para calcular las frecuencias esperadas.

$$p(x) = \frac{\lambda^x \times e^{-\lambda}}{x!} \qquad [5\text{-}4]$$

Al comparar estas frecuencias esperadas con las observadas, podemos examinar la magnitud de la diferencia entre ellas: la desviación absoluta. En la tabla 14-15 enumeramos las frecuencias acumuladas observadas relativas F_0, las frecuencias acumuladas relativas esperadas F_e, y las desviaciones absolutas para $x = 0$ a 22.

Cálculo del estadístico de K-S

Para calcular el estadístico de K-S para este problema, simplemente elija D_n, la desviación absoluta máxima de F_e, alejada de F_o.

Cálculo del estadístico de K-S

Estadístico de K-S

$$D_n = \text{máx} \left| F_e - F_o \right| \qquad [14\text{-}10]$$

En este problema, $D_n = 0.2582$ en $x = 9$.

Cálculo del valor crítico

Una prueba de K-S siempre debe ser una prueba de una cola. Los valores críticos para D_n se han tabulado y pueden encontrarse en la tabla 8 del apéndice. Si buscamos en la fila de $n = 3{,}754$ (el tamaño de muestra) y la columna para un nivel de significancia de 0.01, encontramos que el valor crítico de D_n debe calcularse usando la fórmula:

$$\frac{1.63}{\sqrt{n}} = \frac{1.63}{\sqrt{3{,}754}} = \frac{1.63}{61.27} = 0.0266$$

Nuestra conclusión

El siguiente paso es comparar el valor calculado de D_n con el valor crítico de D_n que se encuentra en la tabla. Si el valor de la tabla para el nivel de significancia elegido es mayor que el valor calculado de D_n, entonces aceptaremos la hipótesis nula. Obviamente, $0.0266 < 0.2582$, así que rechazamos H_0 y llegamos a la conclusión de que una distribución de Poisson con una media de 8.5 *no* es una buena descripción del patrón del uso del transmisor en la central telefónica del condado de Orange.

Tabla 14-15
Frecuencias acumuladas observadas relativas, frecuencias acumuladas observadas esperadas y desviaciones absolutas

Número ocupado	Frecuencia observada	Frecuencia acumulada observada	Frecuencia acumulada observada relativa	Frecuencia acumulada observada esperada	$\lvert F_e - F_o \rvert$ Desviación absoluta
0	0	0	0.0000	0.0002	0.0002
1	5	5	0.0013	0.0019	0.0006
2	14	19	0.0051	0.0093	0.0042
3	24	43	0.0115	0.0301	0.0186
4	57	100	0.0266	0.0744	0.0478
5	111	211	0.0562	0.1496	0.0934
6	197	408	0.1087	0.2562	0.1475
7	278	686	0.1827	0.3856	0.2029
8	378	1,064	0.2834	0.5231	0.2397
9	418	1,482	0.3948	0.6530	0.2582
10	461	1,943	0.5176	0.7634	0.2458
11	433	2,376	0.6329	0.8487	0.2158
12	413	2,789	0.7429	0.9091	0.1662
13	358	3,147	0.8383	0.9486	0.1103
14	219	3,366	0.8966	0.9726	0.0760
15	145	3,511	0.9353	0.9862	0.0509
16	109	3,620	0.9643	0.9934	0.0291
17	57	3,677	0.9795	0.9970	0.0175
18	43	3,720	0.9909	0.9987	0.0078
19	16	3,736	0.9952	0.9995	0.0043
20	7	3,743	0.9971	0.9998	0.0027
21	8	3,751	0.9992	0.9999	0.0007
22	3	3,754	1.0000	1.0000	0.0000

SUGERENCIAS Y SUPOSICIONES

Piense en la prueba de Kolmogorov-Smirnov como otra prueba de *buen ajuste*, igual que la prueba de ji-cuadrada del capítulo 11, excepto que esta vez es más fácil usarla porque no se necesita toda la aritmética del cálculo de la ji-cuadrada. La prueba de K-S sólo encuentra las distribuciones acumuladas relativas tanto para las frecuencias observadas como para las frecuencias esperadas y después prueba qué tan lejanas están. Si la distancia no es significativa, entonces la distribución teórica describe bien a la distribución observada. No pierda de vista que las pruebas de K-S siempre son pruebas de una cola, porque siempre se prueba si las diferencias son mayores que el nivel especificado.

Ejercicios 14-6

Ejercicios de autoevaluación

EA 14-7 La siguiente es una distribución de frecuencias observadas. Usando una distribución normal con $\mu = 6.80$ y $\sigma = 1.24$:

a) Encuentre la probabilidad de caer dentro de cada clase.
b) Del inciso a), calcule la frecuencia esperada de cada categoría.
c) Calcule D_n.
d) Al nivel de significancia de 0.15, ¿parece esta distribución estar bien descrita por la distribución normal sugerida?

Valor de la variable	≤4.009	4.010-5.869	5.870-7.729	7.730-9.589	> 9.590
Frecuencia observada	13	158	437	122	20

Conceptos básicos

■ **14-43** Al nivel de significancia de 0.05, ¿podemos concluir que los siguientes datos provienen de una distribución de Poisson con $\lambda = 3$?

Número de llegadas por día	0	1	2	3	4	5	6 o más
Número de días	6	18	30	24	11	2	9

■ **14-44** La siguiente es una distribución de frecuencias observadas. Utilice la distribución normal con $\mu = 98.6$ y $\sigma = 3.78$.

a) Encuentre la probabilidad de caer en cada clase.
b) Del inciso a), calcule la frecuencia esperada de cada categoría.
c) Calcule D_n.
d) Para un nivel de significancia de 0.10, ¿parece que esta distribución está bien descrita por la distribución normal sugerida?

Valor de la variable	<92.0	92.0-95.99	96.0-99.99	100-103.99	≥104
Frecuencia observada	69	408	842	621	137

■ **14-45** La siguiente es una tabla de frecuencias observadas, junto con las frecuencias que se esperan para una distribución normal.

a) Calcule el estadístico de K-S.
b) ¿Podemos concluir que estos datos en efecto provienen de una distribución normal? Use el nivel de significancia de 0.10.

	Resultados de la prueba				
	51-60	61-70	71-80	81-90	91-100
Frecuencia observada	30	100	440	500	130
Frecuencia esperada	40	170	500	390	100

Aplicaciones

■ **14-46** Kevin Morgan, gerente nacional de ventas de una compañía de electrónica, ha recabado la siguiente estadística de salarios de los ingresos de la fuerza de ventas en su campo. Él tiene tanto las frecuencias observadas como las frecuencias esperadas si la distribución de salarios es normal. Al nivel de significancia de 0.10, ¿puede concluir Kevin que la distribución de los ingresos de la fuerza de ventas es normal?

	Ingresos (miles de dólares)						
	25-30	**31-36**	**37-42**	**43-48**	**49-54**	**55-60**	**61-66**
Frecuencia observada	9	22	25	30	21	12	6
Frecuencia esperada	6	17	32	35	18	13	4

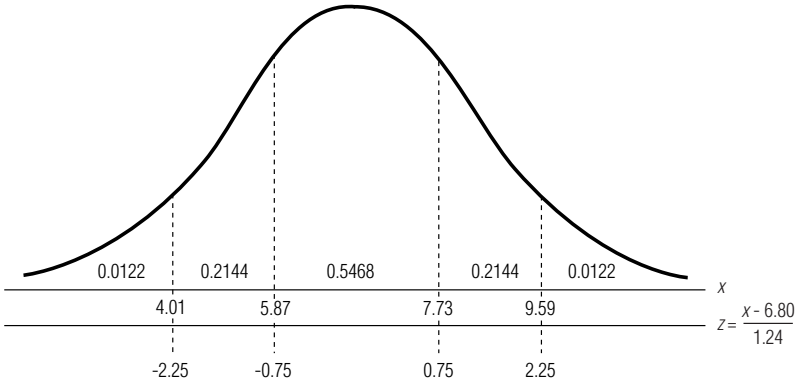
■ **14-47** Randall Nelson, vendedor de la compañía V-Star, tiene que visitar siete cuentas por semana. Se piensa que las ventas del señor Nelson pueden describirse mediante la distribución binomial con probabilidad de venta en cada cuenta de 0.45. Examinando la distribución de frecuencias observadas del número de ventas por semana del señor Nelson, determine si la distribución corresponde en efecto a la distribución sugerida. Use el nivel de significancia de 0.05.

Número de ventas por semana	0	1	2	3	4	5	6	7
Frecuencia del número de ventas	25	32	61	47	39	21	18	12

■ **14-48** Jackie Denn, un administrador de servicios de alimentos para líneas aéreas, ha examinado los registros anteriores de 200 vuelos a través del país, seleccionados aleatoriamente para determinar la frecuencia con que se solicitaron comidas bajas en sodio. El número de vuelos en los que se solicitaron 0, 1, 2, 3, 4 o más comidas bajas en sodio fue 25, 45, 67, 43 y 20, respectivamente. Al nivel de significancia de 0.05, ¿es razonable concluir que estas solicitudes siguen una distribución de Poisson con $\lambda = 1$?

Soluciones a los ejercicios de autoevaluación

EA 14-7 a) Las probabilidades de caer en las cinco clases son las áreas indicadas bajo la curva en la siguiente figura:



b) $n = 13 + 158 + 437 + 122 + 20 = 750$. Entonces, las frecuencias esperadas son $0.0122(750) = 9.15$, $0.2144(750) = 160.80$, $0.5468(750) = 410.1$, 160.80 y 9.15.

c)

f_o	f_o acum.	F_o	F_e	$\lvert F_e - F_o \rvert$
13	13	0.0173	0.0122	0.0051
158	171	0.2280	0.2266	0.0014
437	608	0.8107	0.7734	0.0373 ←
122	730	0.9733	0.9878	0.0145
20	750	1.0000	1.0000	0.0000

d) $D_{tabla} = \dfrac{1.14}{\sqrt{n}} = \dfrac{1.14}{\sqrt{750}} = 0.0416$. $D_n < D_{tabla}$, de manera que se acepta H_0. Los datos están bien descritos por la distribución normal sugerida.

Estadística en el trabajo

Loveland Computers

Caso 14: métodos no paramétricos "Olvidé decirte", dijo Sherrel Wright, la gerente de publicidad, al dirigirse a la oficina, "Margot te estaba buscando, es mejor que la consultes antes de que empieces este proyecto de publicidad".

"¡Necesito ayuda!", anunció Margot con un grito que podía oírse hasta Cheyenne, Wyoming. "Gasté mucho dinero para obtener algunos datos, y ahora que están aquí no sé lo que tengo."

"Bueno, yo tampoco", bromeó Lee, tratando de aligerar los ánimos. "¿Por qué no me dices lo que pasa?"

"Algunos de los modelos de mediano alcance —las PC con chips rápidos y una cantidad razonable de almacena-

miento en disco—, podemos hacerlos en tres presentaciones. Las viejas máquinas estilo AT son del tamaño de una maleta pequeña. A la gente le gustaba la caja grande porque tenía la imagen de una máquina grande y poderosa. Pero en este último año, más o menos, algunas de las estaciones de trabajo más poderosas se han fabricado en un formato de caja de pizza, bastante estrecha y plana. Así que algunas compañías ofrecen ahora las máquinas de mediano alcance en un formato de perfil bajo. En realidad tienen lo mismo en una caja más pequeña que no ocupa tanto espacio de escritorio. Finalmente, algunos competidores ofrecen una configuración de torre. Éste es el viejo estilo AT parado sobre un lado para que se pueda colocar en el piso. Eso elimina la necesidad de espacio de escritorio."

"Así que, ¿qué estilo adoptó por Loveland?", preguntó Lee.

"Francamente, hemos abarcado todos durante las distintas campañas de publicidad. Algunas veces ofrecimos dos de los tres formatos, pero hemos cambiado entre uno y otros intentando adivinar los gustos de los clientes. Uno pensaría que todos preferirían tener la máquina en el piso, pero resulta que la caja de la computadora es un lugar útil para poner el monitor, y la gente que usa muchos discos flexibles no quiere tener que agacharse cada vez que usa la unidad de discos."

"Bueno. Entonces ofrezcan los tres estilos", sonrió Lee ante esta solución simple pero elegante.

"Eso nos sube los costos. Si trabajamos tres estilos, perdemos los descuentos por volumen que obtenemos al trabajar sólo uno. Después tenemos que publicitar tres formatos al tiempo que estoy lanzando nuevos productos de alto perfil e intentando satisfacer la demanda de nuestras máquinas de menor precio. Quisiera recomendar sólo el mejor formato a la gerencia."

"Bueno, yo no tengo una bola de cristal", comenzó a decir Lee.

"No espero que la tengas. Contraté a una compañía de investigación de mercados. Trabajaron con grupos de enfoque en Boulder, New Jersey y Oregon. Había ocho personas en cada grupo, y dos grupos en cada lugar, así que en conjunto obtuve 48 tarjetas de respuestas, y varias horas de discusiones grabadas que no tendrás que ver. Como podrás imaginar, le pedimos a los participantes que clasificaran los tres formatos en términos del estilo que preferirían si fueran a comprar una computadora personal. Después les preguntamos si su primera opción no estuviera disponible, cuál de las otras dos preferirían. Dime cómo vamos a darle sentido a esto para que pueda hacer una recomendación al grupo de planeación de productos."

Preguntas de estudio: ¿Cómo debe organizar los datos Lee y qué pruebas estadísticas son apropiadas? ¿Qué debe hacer Loveland si el análisis de los datos de este pequeño grupo no fuera concluyente?

Ejercicio de base de datos computacional

HH Industries

Todavía sensible respecto a lo que la suposición de Poisson había hecho en el estudio de las llamadas telefónicas, Laurel meditó acerca de las preguntas de la comisión de ventas que recientemente había analizado para Stan. La suposición de poblaciones normales con la misma desviación estándar era crítica para las conclusiones a las que había llegado. "Sólo para asegurarme", pensó, "creo que verificaré mis resultados con algún análisis de suma de rangos".

1. Usando los datos de ventas del capítulo 11, pruebe la suposición de que estas muestras vienen de poblaciones con la misma media. ¿Qué sucedería si se excluye a Mike como antes? Pruebe al nivel de significancia de 0.05.

La siguiente semana, durante el almuerzo, Gary y Laurel discutieron los resultados del estudio de criterios de contratación que acababan de terminar. "Según parece", dijo Gary, "ese estudio es todavía más oportuno de lo que inicialmente pensamos. Parece que Hal tiene planes de establecer otro almacén satélite, esta vez en la región del medio oeste.

Supongo que nuestro negocio es tan sólido que estamos avanzando bastante rápido. Esto seguramente facilitará mi trabajo del personal".

"A propósito", continuó, "¿recuerdas cuando hablamos sobre el estudio de UPS en nuestro almacén de Pennsylvania? Fue poco después de que llegaste, e hiciste un maravilloso análisis sobre nuestro éxito cubrir el área geográfica señalada".

"Oh, oh", bromeó Laurel. "Sé que cuando me alabas, ¡tienes más trabajo sacado de la manga para mí! Sí, recuerdo nuestras conversaciones. ¿Qué hay respecto a ello?"

"Recordarás", prosiguió Gary, "que hice algunos comentarios crípticos acerca de lo que sentí que *debería haber estado*, contra lo que realmente *estaba*, incluido en el estudio. De cualquier forma, esto me ha estado preocupando y quisiera saber tu opinión".

"Continúa", suspiró Laurel, fingiendo irritación. "¡Creo que he creado un monstruo estadístico!"

Gary se rió. "¡Realmente sabes cómo hacer que uno se sienta mal! Ya, en serio, no creo que esto sea mucho problema, sólo tengo curiosidad sobre algo. Cuando UPS hizo ese estudio, sólo tomó en cuenta paquetes que *enviamos*. Resulta que los costos de *recepción* también son bastante significativos. Y, como podrás adivinar, aquí en Florida pagamos una 'penalización' adicional por estar al final de la ruta de en-

vío. La mayor parte de los camiones que nos hacen entregas se van vacíos el estado, porque no hay carga que requiera entregarse de regreso a otras áreas del país."

"Estoy formulando una propuesta que en esencia sugiere que nuestra nueva sucursal en el medio oeste se convierta en nuestro almacén *principal*. Mantendríamos todas nuestras funciones administrativas aquí en Florida; pero, desde el punto de vista de inventarios, actuaríamos solamente como un satélite. ¿Sería posible clasificar todos los estados a los que hacemos envíos y de los que recibimos, sólo para ver si hay alguna correlación allí? Eso me daría alguna información preliminar para mi propuesta.

Laurel se dio cuenta que había estado conteniendo la respiración. "¡Vaya!", atinó a decir. "¡Esa pregunta es mucho más fácil de lo que esperaba! Todo el estudio suena interesante *y* tardado. Pero la respuesta a la pregunta de hoy es sí, eso debe ser relativamente sencillo. Quizá la tendré para finales de esta semana. Y gracias por el aviso. Sé que regresarás con más peticiones de ayuda, ¡y la próxima no será fácil! Voy a empezar a recabar algunos datos que creo pueden ser útiles. No dejes de comunicarme cómo van las cosas."

"De alguna manera pensé que podía contar contigo. Gracias, Laurel", Gary sonrió. "El almuerzo corre por mi cuenta."

2. Usando los datos de los archivos CH14.xxx incluidos en el CD que acompaña al libro, calcule el coeficiente de correlación de rangos entre los estados a los que HH Industries hace envíos y de los que los recibe. Al nivel de significancia de 0.01, ¿podemos concluir que existe una relación entre estos dos factores?

Del libro de texto al mundo real

La estadística en la medicina

Los métodos estadísticos se usan a menudo en la investigación del origen, tratamiento y control de diversas enfermedades. Como que gran parte de los datos de la investigación médica no se ajusta a la distribución normal, los métodos no paramétricos son particularmente útiles. Los doctores Charles H. Kirkpatrick y David W. Alling, aplicaron la prueba de Mann-Whitney de manera ingeniosa para evaluar los resultados de una prueba clínica aleatoria que involucraba el tratamiento de candidiasis oral crónica, una enfermedad caracterizada por infecciones recurrentes de la piel, uñas y membranas mucosas. Los resultados de sus pruebas indicaron que el clotrimazol, que se había empleado con éxito en desórdenes similares, era un tratamiento altamente efectivo para la candidiasis.

La prueba clínica Veinte pacientes que sufrían candidiasis oral persistente fueron admitidos en el estudio y se les asignaron en forma aleatoria tratamientos con pastillas de clotrimazol o placebo. La respuesta de los sujetos al tratamiento se evaluó de dos a siete días después de administrarlo, como se muestra en la tabla MR14-1. Este formato capta dos tipos de resultados y los combina de forma que la mayor de cualesquier dos calificaciones connota el resultado menos favorable; estas calificaciones definen una clasificación ordenada. Los resultados de los tratamientos con clotrimazol y placebo se resumen en la tabla MR14-2. Los 10 pacientes de las pastillas de clotrimazol no presentaron síntomas hacia el quinto día de tratamiento. Esta observación visual fue confirmada por una prueba de Mann-Whitney, que ofreció un fuerte respaldo estadístico.

Tabla MR14-1 **Sistema de calificación para resultados del tratamiento de candidiasis oral crónica**

Calificación	Hallazgos clínicos	Hallazgos de laboratorio
1	Ausente	Negativo
2	Mejorado	Negativo
3	Mejorado	Positivo
4	No mejorado	Positivo

Tabla MR14-2 **Resultado después de 2 a 7 días de tratamiento en 20 pacientes**

	Calificación de resultados				
Grupo de tratamiento	1	2	3	4	Total de pacientes
Clotrimazol	6	3	1	0	10
Placebo	1	0	0	9	10

La enseñanza Los tratamientos exitosos para enfermedades se encuentran sólo por medio de la investigación. En este caso, aunque se sabe que el clotrimazol ocasiona efectos colaterales adversos cuando se administra durante un periodo prolongado, los estudios preliminares que emplean clotrimazol oral con una programación intermitente han mostrado beneficios clínicos. El uso de métodos estadísticos permite a los investigadores clínicos cuantificar resultados de tratamientos médicos, lo que les presta credibilidad a sus hallazgos.

Fuente: C. H. Kirkpatrick y D. W. Alling, "Treatment of Chronic Oral Candidiasis with Clotrimazole Troches: A Controlled Clinical Trial", en *The New England Journal of Medicine* 299(1978): 1201-1203.

Repaso del capítulo

● Términos introducidos en el capítulo 14

Coeficiente de correlación de rangos Medida del grado de asociación entre dos variables que se basa en los rangos o clasificación de las observaciones, no en sus valores numéricos.

Correlación de rango Método para hacer análisis de correlación cuando los datos no están disponibles en forma numérica, pero se tiene información suficiente para clasificar los datos.

Corrida Secuencia de ocurrencias idénticas precedidas y seguidas de ocurrencias diferentes o por ninguna.

Prueba de corridas de una sola muestra Método no paramétrico para determinar la aleatoriedad con que se seleccionaron los elementos de una muestra.

Prueba de Kolmogorov-Smirnov Prueba no paramétrica que no requiere que los datos se agrupen para determinar si existe una diferencia significativa entre la distribución de frecuencias observada y la distribución de frecuencias teórica.

Prueba de Kruskal-Wallis Método no paramétrico para probar si tres o más muestras independientes se extrajeron de poblaciones con la misma distribución. Es una versión no paramétrica de ANOVA, que estudiamos en el capítulo 11.

Prueba de los signos Prueba para la diferencia entre las observaciones por pares, donde los signos + y − sustituyen a los valores cuantitativos.

Pruebas de sumas de rangos Familia de pruebas no paramétricas que usan la información de orden en un conjunto de datos.

Pruebas no paramétricas Técnicas estadísticas que no hacen suposiciones restrictivas respecto a la forma de la distribución de población al realizar una prueba de hipótesis.

Prueba *U* de Mann-Whitney Método no paramétrico usado para determinar si dos muestras independientes se extrajeron de poblaciones con la misma distribución.

Teoría de corridas Teoría desarrollada para permitimos probar la aleatoriedad del orden de las muestras.

● Ecuaciones introducidas en el capítulo 14

■ **14-1**

$$U = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$

Para aplicar la prueba *U* de Mann-Whitney, se necesita esta fórmula para derivar el estadístico *U*, una medida de la diferencia entre las observaciones clasificadas y las dos variables. R_1 es la suma de los rangos de las observaciones de la variable 1; n_1 y n_2 son los números de elementos de las muestras 1 y 2, respectivamente. No es necesario que ambas sean del mismo tamaño.

■ **14-2**

$$\mu_U = \frac{n_1 n_2}{2}$$

Si la hipótesis nula de una prueba *U* de Mann-Whitney es que $n_1 + n_2$ observaciones provienen de poblaciones idénticas, entonces la estadística *U* tiene una distribución de muestreo con una media igual al producto de n_1 y n_2 dividido entre 2.

■ **14-3**

$$\sigma_U = \sqrt{\frac{n_1 n_2(n_1 + n_2 + 1)}{12}}$$

Esta fórmula nos permite derivar el *error estándar del estadístico U* de una prueba *U* de Mann-Whitney.

■ **14-4**

$$U = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2$$

Esta fórmula y la ecuación 14-1 pueden utilizarse de manera intercambiable para derivar el estadístico *U* en la prueba *U* de Mann-Whitney. Para ahorrar tiempo, utilice esta fórmula si el número de observaciones en la muestra 2 es significativamente menor que el número de observaciones en la muestra 1.

■ **14-5**

$$K = \frac{12}{n(n + 1)} \sum \frac{R_j^2}{n_j} - 3(n + 1)$$

La fórmula calcula el estadístico *K* usada en la prueba de Kruskal-Wallis para diferentes medias entre tres o más poblaciones. La distribución de muestreo apropiada para *K* es ji-cuadrada con $k - 1$ grados de libertad, cuando cada muestra contiene al menos cinco observaciones.

■ **14-6**

$$\mu_r = \frac{2n_1 n_2}{n_1 + n_2} + 1$$

Al hacer una prueba de corridas de una sola muestra, use esta fórmula para derivar la media de la distribución muestral del *estadístico r*. Esta *r* es igual al *número de corridas* en la muestra que se está probando.

■ **14-7**
$$\sigma_r = \sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}}$$

Esta fórmula nos permite derivar el *error estándar del estadístico r* en una prueba de corridas de una sola muestra.

■ **14-8**
$$r_s = 1 - \frac{6\Sigma d^2}{n(n^2 - 1)}$$

El *coeficiente de correlación de rangos*, r_s, es una medida de la cercanía de asociación entre dos variables clasificadas.

■ **14-9**
$$\sigma_{r_s} = \frac{1}{\sqrt{n - 1}}$$

Esta fórmula nos permite calcular el *error estándar de* r_s, en una prueba de hipótesis sobre el coeficiente de correlación de rangos.

■ **14-10**
$$D_n = \text{máx } |F_e - F_o|$$

Si comparamos este valor calculado con el valor crítico de D_n en la tabla K-S, podemos probar la bondad de ajuste de distribución.

● Ejercicios de repaso

■ **14-49** Un entrenador de fútbol universitario tiene la teoría de que, en atletismo, el éxito se alimenta de sí mismo. En otras palabras, piensa que ganar un campeonato un año incrementa la motivación del equipo para ganar el siguiente. Expresó esta teoría a un estudiante de estadística, quien le pidió los registros de éxitos y fracasos del equipo de los últimos años. El entrenador le dio una lista, especificando si el equipo había ganado (G) o perdido (P) el campeonato ese año. Los resultados de esta cuenta son:

G, G, G, G, G, G P, G, G, G, G, G, P, G, G, G, G, P, P, G, G, G, G, G, G

a) A un nivel de significancia de 10%, ¿es aleatoria la ocurrencia de éxitos y fracasos?
b) Su respuesta al inciso a), combinada con una inspección ocular de los datos, ¿le dice algo sobre la prueba de corridas de una sola muestra?

■ **14-50** Un pequeño aeropuerto metropolitano abrió una nueva pista, creando una nueva trayectoria de vuelos sobre un área residencial de ingresos altos. Las quejas por ruido excesivo han inundado a las autoridades aeroportuarias hasta el punto de que las dos principales aerolíneas que dan servicio a la ciudad han instalado pantallas acústicas sobre las turbinas de los jets para reducir el ruido y ayudar a aliviar la presión sobre las autoridades. Ambas aerolíneas desean saber si las pantallas han ayudado a reducir el número de quejas levantadas contra el aeropuerto. Si no es así, eliminarán las pantallas porque incrementan el consumo de combustible. Basándose en las siguientes muestras aleatorias de 13 días antes de la instalación de las pantallas y 13 días después, ¿podría decirse, a un nivel de significancia de 0.02, que la instalación de las pantallas ha reducido el número de quejas?

Quejas antes y después de instalar de pantallas acústicas

Antes	27	15	20	24	13	18	30	46	15	29	17	21	18
Después	26	23	19	12	25	9	16	12	28	20	16	14	11

■ **14-51** La American Broadcasting System (ABS) ha invertido una considerable cantidad de dinero en un nuevo programa de televisión, *High Times*, que constituyó la entrada de ABS al mercado de comedias de situación y presentaba la vida despreocupada de un dormitorio universitario. Desafortunadamente, el programa no ha tenido el éxito esperado, y el patrocinador está considerando su cancelación. Para fortalecer los *ratings*, ABS introdujo dormitorios mixtos en la serie. Los siguientes son los resultados de encuestas telefónicas antes y después del cambio. Las encuestas se hicieron en varias áreas metropolitanas importantes, de manera que los resultados son una combinación de las ciudades.

a) Usando una prueba *U*, ¿puede inferir al nivel de significancia de 0.10 que el cambio en el formato de la serie ayudó a los *ratings*?

b) ¿Los resultados de su prueba dicen algo respecto al efecto del sexo en los *ratings* de los programas de televisión?

Proporción de audiencia antes y después del cambio a dormitorios mixtos

Antes	22	18	19	20	31	22	25	19	22	24	18	16	14	28	23	15	16
Después	25	28	18	30	33	25	29	32	19	16	30	33	17	25			

■ **14-52** Los oficiales de personal realizan evaluaciones de preparación general para las unidades militares; la calificación máxima para estas pruebas es 100 puntos. Los oficiales de transporte se quejan de que los evalúan más bajo que a los oficiales de infantería, porque la mayoría de los oficiales de personal vienen de las filas de infantería. A un nivel de significancia de 0.05, pruebe la hipótesis de que no hay diferencia en las calificaciones de ambas unidades durante 10 semanas elegidas al azar.

Tabla de calificaciones

Comando de infantería	72	80	86	90	95	92	88	96	91	82
Comando de transporte	80	79	90	82	81	84	78	74	85	71

La tabla MR12-1, presenta los resultados de las calificaciones del *1992 Business Week* y *U.S. News & World Report* de las escuelas de administración de Estados Unidos. Use esa información para resolver los ejercicios 14-53 y 14-54.

■ **14-53** Considere las 10 mejores escuelas en la clasificación global del *Business Week*. Reajustando las clasificaciones de estudiantes y compañías de reclutamiento para esas 10 escuelas, obtenemos

Escuela	Puntuaciones por estudiantes	Puntuaciones por compañías
Northwestern	2	1
Chicago	7	4
Harvard	8	3
Wharton	9	2
Michigan	6	6
Dartmouth	1	10
Stanford	3	7
Indiana	4	8
Columbia	10	5
North Carolina	5	9

Para $\alpha = 0.10$, ¿difieren las puntuaciones de las compañías y las de los estudiantes?

■ **14-54** Considerando las 20 escuelas, ¿difieren significativamente las puntuaciones de las dos revistas para $\alpha = 0.10$?

■ **14-55** El Comité de Medios y Arbitrios de la Cámara de Representantes de Estados Unidos intentaba evaluar los resultados de bajar los impuestos a personas físicas durante el año anterior. El propósito era estimular la economía, apoyados en la teoría de que con la reducción, el consumidor podría gastar los ahorros de los impuestos. El comité empleó un grupo de investigación de consumidores independiente para seleccionar una muestra de casas y mantener registros de gastos de consumidores tanto antes como después de la puesta en marcha de la legislación. A continuación se muestra una porción de los datos del grupo de investigación:

Programa de gastos del consumidor

Casa	Antes de la legislación	Después de la legislación	Casa	Antes de la legislación	Después de la legislación
1	$ 3,578	$ 4,296	8	22,350	22,500
2	10,856	9,000	9	7,346	7,250
3	7,450	8,200	10	10,345	10,673
4	9,200	9,200	11	5,298	5,349
5	8,760	8,840	12	6,950	7,000
6	4,500	4,620	13	34,782	33,892
7	15,000	14,500	14	12,837	14,297

(*continúa*)

Programa de gastos del consumidor

Casa	Antes de la legislación	Después de la legislación	Casa	Antes de la legislación	Después de la legislación
15	7,926	8,437	24	14,673	13,500
16	5,789	6,006	25	5,003	6,072
17	11,597	12,093	26	10,940	11,398
18	9,612	9,675	27	8,000	9,007
19	3,461	3,740	28	14,256	14,500
20	4,500	4,500	29	4,322	4,258
21	8,341	8,500	30	6,828	7,204
22	7,589	7,609	31	7,549	7,678
23	25,750	24,321	32	8,129	8,125

Para un nivel de significancia del 3%, determine si la política de reducción de impuestos logró los objetivos deseados.

■ **14-56** Muchas compañías de entretenimiento han invertido en parques de diversión vinculados con películas exitosas. La asistencia depende de muchos factores que incluyen el clima. ¿Debe considerarse el clima un evento aleatorio?

■ **14-57** Dos pronosticadores del estado del tiempo en la televisión discutían un día sobre si los años con grandes lluvias tendían a ocurrir en rachas. Uno de ellos dijo que pensaba que había patrones de cantidades de lluvia anual, y que varios años húmedos a menudo eran seguidos por algunos años más secos que el promedio. La otra pronosticadora se mostraba escéptica y dijo que pensaba que la cantidad de lluvias durante años consecutivos era bastante aleatoria. Para investigar la cuestión, decidieron revisar las lluvias anuales abajo (B) o arriba (A) de la mediana de lluvia anual. A continuación se presenta un resumen de sus resultados:

A, A, A, B, B, B, A, B, A, A, B, B, A, B, A, B, A, A, B, B, A, A, A, B, A, A,
A, A, A, B, B, B, A, B, B, B, A, B, A, A, A, B, A, A, A, B, A, B, B, B, A, B, B

Si los pronosticadores prueban a un nivel de significancia del 5%, ¿concluirán que las cantidades anuales de lluvia no ocurren con patrones?

■ **14-58** Anne J. Montgomery, directora administrativa de educación ejecutiva en la Southern University, usa dos tipos de material promocional para anunciar seminarios: cartas personales y folletos. Tiene una gran confianza en que los folletos son el método más efectivo. Ha recolectado datos sobre el número de personas que asistieron a cada uno de los últimos 10 seminarios promovidos por cada método. Al nivel de significancia de 0.15, ¿es acertado lo que cree?

	Número de asistentes									
Carta personal	35	85	90	92	88	46	78	57	85	67
Folleto	42	74	82	87	45	73	89	75	60	94

■ **14-59** La Asociación Nacional para Mejor Publicidad para Niños (NABAC), un grupo de consumidores que promueve mejoras en la televisión para niños, realizó un estudio sobre el efecto de la publicidad los sábados en la mañana. Específicamente, el grupo deseaba saber si la publicidad dirigida a los niños estimulaba un grado significativo de compras, y si existía una correlación positiva entre el tiempo de publicidad en la televisión los sábados en la mañana y las ventas de productos.

La NABAC seleccionó el mercado de cereales de desayuno para niños como un grupo de muestra. Seleccionó productos cuyo mensaje publicitario estuviera dirigido totalmente a los niños. Los resultados del estudio se presentan a continuación. (El cereal de ventas más altas tiene rango l.)

Comparación de tiempo de publicidad por televisión y ventas de productos

Producto	Tiempo de publicidad en minutos	Rango de ventas
Captain Grumbles	0.50	10
Obnoxious Berries	3.00	1
Fruity Hoops	1.25	9
OO La Granola	2.00	5

(*continúa*)

Comparación de tiempo de publicidad por televisión y ventas de productos

Producto	Tiempo de publicidad en minutos	Rango de ventas
Sweet Tweets	3.50	2
Chocolate Chumps	1.00	11
Sugar Spots	4.00	3
County Cavity	2.50	8
Crunchy Munchies	1.75	6
Karamel Kooks	2.25	4
Flakey Flakes	1.50	7

¿Puede el grupo concluir que existe una correlación de rangos positiva entre el tiempo de publicidad matutina en televisión y el volumen de ventas de los cereales para niños? Pruebe a un nivel de significancia del 5%.

Los fondos mutuos proporcionan la oportunidad de participar en el mercado de valores a los pequeños inversionistas, quienes pueden seleccionar entre más de 3,000 fondos, buscando aquellos que les proporcionen seguridad a sus ahorros *y* los mayores rendimientos totales (los dividendos ganados por el fondo más los incrementos en el precio de las acciones propiedad del fondo). Los fondos exitosos atraen más inversiones, de modo que la cantidad de dinero que se maneja puede ascender a varios miles de millones de dólares. Algunos observadores del mercado consideran un tamaño grande de fondo como una desventaja, porque puede tener tantas acciones de una compañía particular que un intento de liquidar una posición puede bajar el precio. La tabla MR14-3 presenta información del tamaño y los rendimientos de los 20 fondos mutuos más grandes a finales de septiembre de 1992. Use esa información para hacer los ejercicios 14-60 y 14-61.

14-60 a) ¿Existe una relación significativa entre el tamaño del fondo y los rendimientos totales promedio ganados en los últimos 5 años?

b) ¿Existe una relación significativa entre los rendimientos totales de 1992 y los rendimientos de los últimos 5 años?

Tabla MR 14-3

Los 20 fondos mutuos más grandes, septiembre de 1992

Fondo	Tamaño (miles de millones de dólares)	Rendimientos totales anualizados (%) 1992	Prom. de 5 años
Fidelity Magellan	21.05	9.51	11.24
Investment Co. of America	14.03	11.08	9.50
Washington Mutual	9.48	11.35	8.99
Windsor	8.23	9.53	7.00
Vanguard Index 500	5.77	10.87	8.73
Income Fund of America	5.64	16.33	11.57
Fidelity Puritan	5.62	15.11	9.38
Wellington	5.10	11.44	9.34
AIM Weingarten	4.98	5.77	11.07
Windsor II	4.80	14.71	9.59
Dean Witter Dividend	4.67	11.42	10.03
Janus	4.66	8.55	14.70
Fidelity Equity-Income	4.65	12.45	7.29
American Mutual	4.60	11.59	9.06
Twentieth Century Select	4.47	2.02	6.25
Twentieth Century Growth	4.40	10.84	9.52
Fidelity Growth & Income	4.29	10.51	11.80
Templeton World	4.02	7.00	5.47
Twentieth Century Ultra	4.01	1.24	14.55
Pioneer II	3.97	9.92	4.78

Fuente: Standard & Poor's/Lipper Mutual Fund Profiles, *6(4) (noviembre de 1992).*

14-61 Dos administradores de fondos grandes, Fidelity y Twentieth Century, tienen varios registros en la lista de la tabla MR14-3. ¿El rendimiento de los fondos ofrecidos por estas dos compañías proviene de una distribución diferente del rendimiento de los restantes 13 fondos? Pruebe esto para los rendimientos totales de 1992 y para los rendimientos totales promedio de 5 años.

■ **14-62** *American Motoring Magazine* probó recientemente la efectividad de frenado de dos materiales de discos de freno. Los datos que representan las distancias de frenado para ambos tipos de materiales se presentan a continuación. Al nivel de significancia de 0.05, pruebe la hipótesis de que no hay diferencia en la efectividad de los materiales.

	Distancia de frenado (pies)									
Unión de grafito	110	120	130	110	100	105	110	130	145	125
Bronce aglomerado	100	110	135	105	105	100	100	115	135	120

■ **14-63** Como parte de una encuesta sobre la calidad de restaurantes, una revista local pidió a los residentes del área que clasificaran dos restaurantes especializados en cortes de carne. Los sujetos debían valorar características como calidad de la comida, atmósfera, servicio y precio en una escala de 1 a 10. Después de recolectar los datos, uno de los propietarios de los restaurantes propuso que se realizaran varias pruebas estadísticas. Específicamente mencionó que le gustaría ver una media y una desviación estándar para las respuestas a cada pregunta sobre cada restaurante, con el fin de ver cuál había calificado mejor. Varios empleados de la revista refutaron estas sugerencias, observando que la calidad de los datos de entrada no justificaría un análisis estadístico detallado. Argumentaron que lo importante era como clasificaban los residentes a los dos restaurantes. Evalúe los argumentos presentados por el propietario del restaurante y por los empleados de la revista.

■ **14-64** Se pidió a algunos estudiantes de administración del último año entrevistados por la compañía de seguros Ohio no comentar sus entrevistas con otros estudiantes de la escuela hasta que se hubiera ido la reclutadora. Sin embargo, ella sospechó que los últimos solicitantes sabían más acerca de qué buscaba. ¿Eran correctas sus sospechas? Para descubrirlo, clasifique las puntuaciones de la entrevista que obtuvieron los sujetos dados en la tabla. Después pruebe la significancia del coeficiente de correlación de rangos entre las puntuaciones y el número de entrevista. Use el nivel de significancia de 0.02.

Número de entrevista	Puntuación	Número de entrevista	Puntuación	Número de entrevista	Puntuación	Número de entrevista	Puntuación
1	63	6	57	11	77	16	70
2	59	7	76	12	61	17	75
3	50	8	81	13	53	18	90
4	60	9	58	14	74	19	80
5	66	10	65	15	82	20	89

■ **14-65** Hace más de tres años, la Administración de Salud y Seguridad Ocupacional de Estados Unidos (OSHA) requirió la implantación de varias medidas de seguridad en la planta de aluminio de Northbridge. Ahora OSHA desea saber si los cambios dieron como resultado menos accidentes en la planta. Se recolectaron estos datos:

	Accidentes en la planta de Northbridge											
	Ene.	Feb.	Mar.	Abr.	May.	Jun.	Jul.	Ago.	Sep.	Oct.	Nov.	Dic.
1992	5	3	4	2	6	4	3	3	2	4	5	3
1993	4	4	3	3	3	4	0	5	4	2	0	1
1994	3	2	1	1	0	2	4	3	2	1	1	2
1995	2	1	0	0	1	2						

a) Determine la mediana del número de accidentes por mes. Si las medidas de seguridad han sido efectivas, debemos encontrar que los primeros meses tienen un valor más alto que la mediana y los últimos más bajo. En consecuencia, habrá un pequeño número de corridas arriba y abajo de la mediana. Realice una prueba al nivel de significancia de 0.03 para ver si los accidentes están distribuidos aleatoriamente.

b) ¿Qué puede concluir respecto a la efectividad de las medidas de seguridad?

■ **14-66** Un servicio de ambulancias del condado calcula que por cada municipio que atiende, durante un turno dado de seis horas, existe un 35% de probabilidad de recibir al menos una llamada de asistencia. La siguiente es una muestra aleatoria de 90 días:

Número de turnos durante los cuales se recibieron llamadas	0	1	2	3	4
Número de días	5	35	30	13	7

Al nivel de significancia de 0.05, ¿siguen estas llamadas de asistencia una distribución binomial?

La empresa Overseas Shipholding Group, Inc., tiene interés en comprar 64 buques, clasificados como cargueros, transportadores de productos petroleros (PPC) o buques-tanque. Los tipos de buques, registro y los años en los que fueron construidos se dan en la tabla MR14-4. Use esta información para hacer los ejercicios 14-67 y 14-68.

14-67 ¿Tienen la misma antigüedad las flotas de Estados Unidos y las extranjeras? Es decir, ¿puede concluir que los datos de años de construcción para las dos flotas provienen de la misma distribución?

14-68 ¿Son las distribuciones de antigüedad de los tres tipos de transportadores significativamente diferentes?

■ **14-69** Jim Bailey, propietario de Crow's Nest Marina, cree que el número de horas de trabajo de un motor de bote en agua salada, y no la edad del bote, es el mejor pronosticador de fallas del motor. Su gerente de servicio ha recabado datos de sus registros de reparación de motores descompuestos. Al nivel de significancia de 0.05, ¿está Jim en lo correcto?

Motor	Horas en agua salada	Antigüedad del motor (años)	Costo de reparación (dólares)
1	300	4	625
2	150	6	350
3	200	3	390
4	250	6	530
5	100	4	200
6	400	5	1,000
7	275	6	550
8	350	6	800
9	325	3	700
10	375	2	600

Tabla MR14-4
Datos de la flota del Overseas Shipholding Group

Tipo	Registro	Construcción	Tipo	Registro	Construcción	Tipo	Registro	Construcción
De carga	EE.UU.	1978	De carga	No EE.UU.	1973	PPC	No EE.UU.	1981
De carga	EE.UU.	1978	De carga	No EE.UU.	1981	PPC	No EE.UU.	1981
PPC	EE.UU.	1983	De carga	No EE.UU.	1983	PPC	No EE.UU.	1982
PPC	EE.UU.	1982	De carga	No EE.UU.	1983	Tanques	No EE.UU.	1973
PPC	EE.UU.	1969	De carga	No EE.UU.	1989	Tanques	No EE.UU.	1975
PPC	EE.UU.	1968	De carga	No EE.UU.	1989	Tanques	No EE.UU.	1974
PPC	EE.UU.	1968	De carga	No EE.UU.	1980	Tanques	No EE.UU.	1974
Tanques	EE.UU.	1974	De carga	No EE.UU.	1980	Tanques	No EE.UU.	1989
Tanques	EE.UU.	1973	De carga	No EE.UU.	1977	Tanques	No EE.UU.	1990
Tanques	EE.UU.	1977	De carga	No EE.UU.	1975	Tanques	No EE.UU.	1972
Tanques	EE.UU.	1977	De carga	No EE.UU.	1975	Tanques	No EE.UU.	1989
Tanques	EE.UU.	1978	De carga	No EE.UU.	1985	Tanques	No EE.UU.	1989
Tanques	EE.UU.	1977	De carga	No EE.UU.	1985	Tanques	No EE.UU.	1976
Tanques	EE.UU.	1971	PPC	No EE.UU.	1986	Tanques	No EE.UU.	1975
Tanques	EE.UU.	1970	PPC	No EE.UU.	1986	Tanques	No EE.UU.	1975
De carga	No EE.UU.	1982	PPC	No EE.UU.	1986	Tanques	No EE.UU.	1986
De carga	No EE.UU.	1982	PPC	No EE.UU.	1987	Tanques	No EE.UU.	1987
De carga	No EE.UU.	1975	PPC	No EE.UU.	1989	Tanques	No EE.UU.	1980
De carga	No EE.UU.	1975	PPC	No EE.UU.	1988	Tanques	No EE.UU.	1981
De carga	No EE.UU.	1990	PPC	No EE.UU.	1989	Tanques	No EE.UU.	1979
De carga	No EE.UU.	1990	PPC	No EE.UU.	1989			
De carga	No EE.UU.	1973	PPC	No EE.UU.	1979			

Fuente: Overseas Shipholding Group, Inc., *1991 Annual Report*.

■ **14-70** SavEnergy, un grupo activista internacional preocupado por el dominio total de las áreas occidentales en el uso de la energía, ha afirmado que el tamaño de la población y el consumo de energía tienen una correlación negativa. Sus oponentes afirman que no existe correlación. Usando los siguientes datos, pruebe la hipótesis de que no existe una correlación de rangos entre el número de habitantes y el consumo de energía, contra la afirmación de correlación negativa de SavEnergy. Use un nivel de significancia de 0.10.

	Habitantes en 1989 (millones)	Consumo total de energía (10^{15} joules)
Estados Unidos	249	68
Latinoamérica	438	16
África	646	11
Europa	499	65
Unión Soviética	289	54
India	835	9
China	1,100	24

■ **14-71** Los accidentes en carretera ocasionaron la muerte de más de 75,000 ocupantes de autos de pasajeros entre 1995 y 1996. Usando esa sombría estadística como punto de inicio, los investigadores del Instituto de Seguridad en Carreteras calcularon las tasas de decesos para las 103 series de vehículos más vendidos. Los vehículos fueron clasificados como furgonetas y *vans*, autos de cuatro puertas, autos de dos puertas o autos deportivos o especializados. Otra estratificación en cada categoría etiquetó a los vehículos como grandes, medianos o pequeños. Considerando las tasas (decesos por 10,000 vehículos registrados) para autos de cuatro puertas, las cifras son las siguientes:

Grandes	1.2	1.3	1.4	1.5	1.5	1.5	1.6	1.8		
Medianos	1.1	1.2	1.2	1.2	1.3	1.3	1.3	1.3	1.4	1.4
	1.5	1.6	1.6	1.6	1.7	1.7	1.8	1.9	2.0	2.3
	2.3	2.4	2.5	2.6	2.9					
Pequeños	1.1	1.5	1.6	1.7	1.8	2.0	2.0	2.0	2.3	2.5
	2.6	2.8	3.2	4.1						

Use la prueba de Kruskal-Wallis para probar si las tres medias de población son, en efecto, iguales. Pruebe al nivel de significancia de 0.05.

■ **14-72** El año 1996 fue particularmente malo en cuanto a lesiones de jugadores de béisbol profesional. Para los siguientes datos, ¿una prueba de signo para datos por pares indica que los jugadores de la Liga Americana sufrieron significativamente más lesiones que sus análogos de la Liga Nacional? Use un nivel de significancia de 0.05.

Ubicación de la lesión	LA	LN	Ubicación de la lesión	LA	LN
Hombro	46	22	Espalda	10	7
Cuello	3	0	Muñeca	10	2
Costillas	7	5	Cadera	1	1
Codo	21	19	Mano	6	4
Dedo de la mano	7	5	Tobillo	6	4
Muslo	17	14	Pie	1	4
Ingle	7	3	Dedo del pie	0	1
Rodilla	16	18	Otros	10	4

■ **14-73** La investigación continua sobre los patrones de clima que pueden correlacionarse con las manchas solares se han enfocado en la temperatura polar (la temperatura promedio en la estratosfera arriba del Polo Norte) durante periodos en que soplan ciertos vientos ecuatoriales. Cuando estos vientos vienen del oeste, la temperatura polar parece subir y bajar con la actividad solar. Cuando los vientos son del este, la temperatura parece hacer lo opuesto de lo que el sol hace. A partir de los datos, calcule los coeficiente de correlación de rango entre estas variables y pruebe, al nivel de significancia de 0.05, si las relaciones hipotéticas se mantienen (correlación positiva para vientos del oeste, correlación negativa para vientos del este).

Actividad solar	Temperatura polar (°F) Vientos del este	Vientos del oeste
230	−85	−76
160	−97	−86
95	−88	−100
75	−85	−110
100	−90	−108
165	−96	−85
155	−91	−70
120	−76	−100
75	−80	−110
65	−86	−112
125	−90	−99
195	−104	−91
190	−95	−93
125	−99	−99
75	−73	−103

■ **14-74** El 31 de diciembre de 1992, las autoridades de la autopista de Nueva Jersey emitieron una convocatoria para redimir algunos de sus bonos de $5,000. Los bonos eran parte de una emisión de $45,000,000 en 9,000 bonos, con una tasa de interés del 4.85% y originalmente pagadera en 2001. Las autoridades tenían efectivo a la mano o encontraron una fuente de dinero menos costosa, y exigieron el cobro de los bonos para reducir sus gastos de intereses. Los bonos seleccionados para su rescate se numeraron:

5	8	16	21	183	646	1018	1044	1045	1233
1616	1794	1908	2069	2072	2172	2922	2927	2930	3044
3925	4225	4283	4303	4312	4316	4347	4380	5189	6266
6281	6291	7107	7114	7128	7145	7230	7394	8162	8172
8179	8180								

Fuente: The Wall Street Journal *(7 de enero de 1993): C20.*

a) Suponiendo que los bonos se seleccionaron aleatoriamente para su rescate, ¿cuántos esperaría ver con números entre 1 y 1500, 1501 y 3000, . . . , 7501 y 9000?
b) Use la prueba de bondad de ajuste ji-cuadrada, estudiada en el capítulo 11, para ver si es razonable concluir que los bonos cobrados se eligieron aleatoriamente.
c) Use la prueba de Kolmogorov-Smirnov para ver si la selección aleatoria de los bonos cobrados es una hipótesis razonable.
d) Compare sus resultados de los incisos b) y c).

■ **14-75** Los analistas "técnicos" de acciones examinan cuidadosamente los patrones comerciales de los precios de las acciones y, en el límite, ignoran las noticias "fundamentales" del desempaño y proyectos de una compañía (o industria). Un analista observó cada media hora de actividades bursátiles el índice industrial Dow Jones desde enero de 1987 hasta abril de 1991 y registró cuando el promedio subía (S) o bajaba (B) en comparación con la media hora anterior. ¿Es la siguiente secuencia de medias horas en que sube o baja el promedio durante una semana un patrón aleatorio?

Lunes	BSBBSSSSBSSSSS
Martes	SSBBBBSBBBSSBS
Miércoles	SSBSSSSSBBSBSB
Jueves	SBBSSSSSBBSSSB
Viernes	SSBSSSSSBBBSSS

Fuente: Yale Hirsch, The 1993 Stock Trader's Almanac *(Old Tappan, NJ: The Hirsch Organization) 1992, pág. 117).*

■ **14-76** Los gerentes de negocios con operaciones de servicios deben manejar las horas pico en las que llegan muchos clientes al mismo tiempo. El gerente del quiosco de información en un centro comercial de las afue-

ras de la ciudad recolectó los siguientes datos de las llegadas por minuto entre las 7:10 y 8:00 los jueves, noche en que cierra tarde el centro comercial.

Número de llegadas	1	2	3	4	5	6	7	8	9	10	11
Frecuencia	5	3	2	6	6	2	6	10	4	4	2

Pruebe si una distribución de Poisson con una media de seis describe adecuadamente estos datos, use el nivel de significancia de 0.05.

■ **14-77** Los resultados de la primera carrera de 10 km de la Asociación Atlética de Carolina mostraron el siguiente orden de finalistas hombres y mujeres:

H H H H H H H H H H H H M H H H H H H M H H H H
H M H H H H H H H H M H M H H H H M H H H H M H
H M H H H H H H M H H M H H H M M M M H M H M M H
M H H H M H M M H M M M M H H H M H H

¿La llegada a la meta de las mujeres fue aleatoria? Use el nivel de significancia de 0.20.

■ **14-78** Se dio a varios grupos una lista de 30 actividades y avances tecnológicos y se les pidió que los clasificaran considerando el riesgo de morir como consecuencia de cada uno. Los resultados se muestran en la siguiente tabla. Calcule el coeficiente de correlación de rangos de cada grupo relativo a la clasificación de los expertos. ¿Qué grupo pareció tener la percepción más exacta de los riesgos implicados?

A = Expertos
B = Liga de Mujeres Votantes
C = Estudiantes universitarios
D = Miembros del Club Cívico

Riesgo	A	B	C	D
Vehículos de motor	1	2	5	3
Fumar	2	4	3	4
Bebidas alcohólicas	3	6	7	5
Armas de fuego	4	3	2	1
Cirugía	5	10	11	9
Motocicletas	6	5	6	2
Rayos X	7	22	17	24
Pesticidas	8	9	4	15
Energía eléctrica (no nuclear)	9	18	19	19
Natación	10	19	30	17
Anticonceptivos	11	20	9	22
Aviación general (privada)	12	7	15	11
Grandes construcciones	13	12	14	13
Preservadores de comida	14	25	12	28
Bicicletas	15	16	24	14
Aviación comercial	16	17	16	18
Trabajo policial	17	8	8	7
Trabajo de bomberos	18	11	10	6
Transporte ferroviario	19	24	23	20
Energía nuclear	20	1	1	8
Colorantes de comida	21	26	20	30
Aparatos domésticos	22	29	27	27
Caza	23	13	18	10
Antibióticos prescritos	24	28	21	26
Vacunaciones	25	30	29	29
Aerosoles	26	14	13	23
Fútbol de bachillerato y univ.	27	23	26	21
Podadoras eléctricas	28	27	28	25
Alpinismo	29	15	22	12
Esquí	30	21	25	16

■ **14-79** Al probar un nuevo medicamento para la fiebre del heno, los investigadores midieron la incidencia de efectos secundarios adversos de la droga, administrándola a un gran número de pacientes y evaluándolos contra un grupo de control. Se registraron los porcentajes de pacientes que reportaron 13 tipos de efectos secundarios. Usando una prueba de los signos para datos por pares, ¿puede determinar si algún grupo como un todo experimentó más efectos adversos? Use el nivel de significancia de 0.10.

Efectos secundarios	Droga	Control
A	9.0	18.1
B	6.3	3.8
C	2.9	5.8
D	1.4	1.0
E	0.9	0.6
F	0.9	0.2
G	0.6	0.0
H	4.6	2.7
I	2.3	3.5
J	0.9	0.5
K	0.5	0.5
L	0.0	0.2
M	1.0	1.4

■ **14-80** La banca comercial ha experimentado una consolidación sustancial conforme los grandes bancos nacionales han adquirido a los bancos más pequeños, y algunos medianos se han fusionado. La siguiente tabla da la clasificación de los 10 bancos más importantes (por el total de activos) en 1995, y la clasificación que los miembros del grupo tenían entre ellos en 1985. Para un nivel de significancia del 10%, ¿es cierto que "el grande se hace más grande", o la competencia entre los bancos ha llevado a cambios significativos en la clasificación?

Banco	Rango 1995	Rango 1985
Citicorp	1	1
BankAmerica	2	2
NationsBank	3	8
Chemical	4	5
J P. Morgan	5	4
First Union	6	10
Chase Manhattan	7	3
First Chicago	8	7
Bankers Trust	9	6
Banc One	10	9

Fuente: 'Creating New Giants", The Wall Street Journal *(14 de julio de 1995): A4.*

5.5. Resultados y discusión

5.5.1. Interpretación de los resultados en relación con los objetivos e hipótesis – (Interpretación y discusión de los resultados en relación con los objetivos, hipótesis y antecedentes)

- Autor principal: **Roberto Hernández Sampieri**
- Libro recomendado: *Metodología de la Investigación*
- Autor de apoyo: **John W. Creswell**
- Obra: *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*

5.5.2. Elaboración de conclusiones fundamentadas y su relación con las hipótesis

- **Autor principal:** César Augusto Bernal Torres
- **Libro recomendado:** *Metodología de la investigación*
- **Autor de apoyo:** Roberto Hernández Sampieri
- **Obra:** *Metodología de la investigación*

5.5.1. Interpretación de los resultados en relación con los objetivos e hipótesis

Para el desarrollo del subtema **Interpretación y discusión de los resultados en relación con los objetivos, hipótesis y antecedentes**, se consideran aquellos contenidos de *Metodología de la Investigación* que permiten comprender el análisis e interpretación de los resultados obtenidos, su relación con las hipótesis planteadas y la preparación de los resultados para su presentación. Asimismo, se considera la revisión de la literatura como elemento de referencia para relacionar los hallazgos de la investigación con los antecedentes y conocimientos existentes sobre el problema estudiado.

N.°	Libros	Capítulo	Página	Contenido
1	*Metodología de la Investigación* — Roberto Hernández Sampieri	**Capítulo 4. Desarrollo de la perspectiva teórica: revisión de la literatura y construcción del marco teórico**	**p. 68**	**¿Qué nos puede revelar la revisión de la literatura?**
2	*Metodología de la Investigación* — Roberto Hernández Sampieri	**Capítulo 4. Desarrollo de la perspectiva teórica: revisión de la literatura y construcción del marco teórico**	**p. 68**	**¿Qué información o contenido se extrae de las referencias?**
3	*Metodología de la Investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**p. 288**	**¿Cómo se interpretan las medidas de tendencia central y de la variabilidad?**
4	*Metodología de la Investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**p. 299**	**Paso 5: analizar mediante pruebas estadísticas las hipótesis planteadas (análisis estadístico inferencial).**
5	*Metodología de la Investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**p. 299**	**¿En qué consiste la prueba de hipótesis?**
6	*Metodología de la Investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**p. 326**	**Paso 6: realizar análisis adicionales.**

7	*Metodología de la Investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**p. 327**	**Paso 7: preparar los resultados para presentarlos.**
8	*Metodología de la Investigación* — Roberto Hernández Sampieri	**Capítulo 11. El reporte de resultados del proceso cuantitativo**	**p. 340**	**¿Qué apartados o secciones contiene un reporte de investigación o un reporte de resultados en un contexto académico?**

5.5.1. Interpretación y discusión de los resultados en relación con los objetivos, hipótesis y antecedentes

Para el desarrollo del subtema **Interpretación y discusión de los resultados en relación con los objetivos, hipótesis y antecedentes**, se consideran aquellos contenidos relacionados con la interpretación de los resultados obtenidos en una investigación, su vinculación con las preguntas de investigación y las hipótesis, la presentación de los resultados y la valoración de su significancia estadística y práctica. Asimismo, se considera la discusión de los hallazgos en relación con estudios previos, identificando coincidencias, diferencias, posibles aportes, implicaciones, limitaciones y futuras líneas de investigación. Estos contenidos se desarrollan principalmente en el **Capítulo 8. Quantitative Methods** de Creswell y Creswell.

N.°	Libros	Capítulo	Página	Contenido
1	*Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* — John W. Creswell y J. David Creswell (2023)	**Capítulo 8. Quantitative Methods**	**pp. 170–172**	**Interpreting Results and Writing a Discussion Section:** interpretación de los resultados en relación con las preguntas de investigación, las hipótesis y el significado general de los resultados.
2	*Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* — John W. Creswell y J. David Creswell (2023)	**Capítulo 8. Quantitative Methods**	**p. 170**	**Reporting Results:** presentación de los resultados y explicación de cómo los hallazgos responden a la pregunta de investigación o a la hipótesis.
3	*Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* — John W. Creswell y J. David Creswell (2023)	**Capítulo 8. Quantitative Methods**	**pp. 169–172**	**Statistical Significance Testing:** interpretación de la significancia estadística de los resultados y valoración de la evidencia para las preguntas de investigación y las hipótesis.
4	*Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* — John W. Creswell y J. David Creswell (2023)	**Capítulo 8. Quantitative Methods**	**p. 172**	**Practical Evidence:** consideración de los intervalos de confianza y del tamaño del efecto para valorar la importancia práctica de los resultados.
5	*Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* — John W. Creswell y J. David Creswell (2023)	**Capítulo 8. Quantitative Methods**	**p. 172**	**Context of Previous Studies:** discusión de los principales hallazgos en el contexto de estudios publicados anteriormente, considerando si los resultados son consistentes con investigaciones previas, las refutan o las amplían.

6	*Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* — John W. Creswell y J. David Creswell (2023)	**Capítulo 8. Quantitative Methods**	**p. 172**	**Context of Previous Studies:** análisis de las implicaciones de los hallazgos para la práctica, la investigación futura y la teoría, así como reconocimiento de limitaciones y posibles explicaciones alternativas.
7	*Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* — John W. Creswell y J. David Creswell (2023)	**Capítulo 8. Quantitative Methods**	**pp. 185–186**	**Interpreting Results and Writing a Discussion Section:** interpretación y discusión de resultados en investigaciones experimentales, en relación con las preguntas de investigación, hipótesis y literatura previa.
8	*Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* — John W. Creswell y J. David Creswell (2023)	**Capítulo 8. Quantitative Methods**	**pp. 169–172**	**Statistical Tests in Survey Research:** utilización e interpretación de pruebas estadísticas para responder preguntas de investigación y contrastar hipótesis, incluyendo pruebas *t*, ANOVA, ANCOVA, chi-cuadrado, correlación de Pearson y regresión múltiple.
9	*Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* — John W. Creswell y J. David Creswell (2023)	**Capítulo 9. Qualitative Methods**	**pp. 210–212**	**Steps in the Analytic Process — Developing a Story Line Interpretation / Representing and Interpreting the Data:** interpretación de datos cualitativos mediante temas, patrones y significados derivados del análisis.
10	*Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* — John W. Creswell y J. David Creswell (2023)	**Capítulo 10. Mixed Methods Procedures**	**pp. 234–254**	**Integration, Joint Display and Metainferences:** integración de resultados cuantitativos y cualitativos para generar una interpretación conjunta de los hallazgos en investigaciones de métodos mixtos.

Aun en el caso de que desechemos los estudios previos, éstos nos orientarán sobre lo que queremos y lo que no queremos para nuestra investigación.

3. Amplía el horizonte del estudio o guía al investigador para que se centre en su problema y evite desviaciones del planteamiento original.
4. Documenta la necesidad de realizar el estudio.
5. Conduce al establecimiento de hipótesis o afirmaciones que más tarde habrán de someterse a prueba en la realidad, o nos ayuda a no establecerlas por razones bien fundamentadas.
6. Inspira nuevas líneas y áreas de investigación (Race, 2010 y Yurén Camarena, 2000).
7. Provee de un marco de referencia para interpretar los resultados del estudio. Aunque podemos no estar de acuerdo con dicho marco o no utilizarlo para explicar nuestros resultados, es un punto de referencia.

Ejemplo

Sobre una investigación sin sentido por no contar con perspectiva teórica

Si intentamos probar que determinado tipo de personalidad incrementa la posibilidad de que un individuo sea líder, al revisar los estudios sobre liderazgo en la literatura respectiva nos daríamos cuenta de que tal investigación carece de sentido, pues se ha demostrado con amplitud que el liderazgo es más bien producto de la interacción entre tres elementos: características del líder, características de los seguidores (miembros del grupo) y la situación en particular. Por ello, poseer ciertas características de personalidad no está relacionado necesariamente con el surgimiento de un líder en un grupo (no todos los "grandes líderes históricos" eran extravertidos, por ejemplo).

¿Qué etapas comprende el desarrollo de la perspectiva teórica?

Tal desarrollo usualmente comprende dos etapas:

- La revisión analítica de la literatura correspondiente.
- La construcción del marco teórico, lo que puede implicar la adopción de una teoría.

¿En qué consiste la revisión analítica de la literatura? 1

La **revisión de la literatura** implica *detectar, consultar y obtener la bibliografía* (referencias) y otros materiales que sean útiles para los propósitos del estudio, de donde se tiene que *extraer* y *recopilar* la información relevante y necesaria para enmarcar nuestro problema de investigación. Esta revisión debe ser *selectiva*, puesto que cada año se publican en el mundo miles de artículos en revistas académicas y periódicos, libros y otras clases de materiales sobre las diferentes áreas del conocimiento. Si al revisar la literatura nos encontramos con que en el área de interés hay 5 000 posibles referencias, es evidente que se requiere seleccionar sólo las más importantes y recientes, y que además estén directamente vinculadas con nuestro planteamiento del problema de investigación. En ocasiones, revisamos referencias de estudios tanto cuantitativos como cualitativos, sin importar nuestro enfoque, porque se relacionan de manera estrecha con nuestros objetivos y preguntas.

Revisión de la literatura Paso de investigación que consiste en detectar, consultar y obtener la bibliografía y otros materiales útiles para los propósitos del estudio, de los cuales se extrae y recopila información relevante y necesaria para el problema de investigación.

A continuación comentamos los pasos que se siguen para revisar la literatura.

Inicio de la revisión de la literatura 2

La revisión de la literatura puede iniciarse directamente con el acopio de las referencias o fuentes primarias,[1] situación que ocurre cuando el investigador conoce su localización, se encuentra muy

[1] Las *referencias* o *fuentes primarias* proporcionan datos de primera mano, pues se trata de documentos que incluyen los resultados de los estudios correspondientes. Ejemplos de fuentes primarias son: libros, antologías, artículos de publicaciones periódicas, monografías, tesis y disertaciones, documentos oficiales, reportes de asociaciones, trabajos presentados en conferencias o seminarios, artículos periodísticos, testimonios de expertos, documentales, videocintas en diferentes formatos, foros y páginas en internet, etcétera.

familiarizado con el campo de estudio y tiene acceso a ellas (puede utilizar material de bibliotecas, filmotecas, hemerotecas y bancos de información). Sin embargo, frecuentemente no sucede así (no se es experto en el tema o se dispone de recursos limitados).

Por ello, *es recomendable iniciar la revisión de la literatura consultando a uno o varios especialistas en el tema* (algún profesor, por ejemplo) y buscando en internet fuentes primarias en centros o sistemas de información y bases de referencias y datos.

Para ello, necesitamos elegir las "palabras clave", "descriptores" o "términos de búsqueda", los cuales deben ser distintivos del problema de estudio y se extraen de la idea o tema y del planteamiento del problema. Para este último se requiere hacer algunas lecturas preliminares que lo afinen y completen. Los expertos también nos pueden ayudar a seleccionar tales palabras.

Si los términos son vagos y generales obtendremos una consulta con muchas referencias e información que *no* es pertinente para nuestro planteamiento. En este sentido, las bases de referencias funcionan como los motores de búsqueda (Google, Yahoo, Bing, etcétera).

Por ejemplo, si hacemos una consulta con palabras como "escuela", "educación", "comunicación", "empresas" o "personalidad", aparecerán miles de referencias y nos "perderemos en un mundo de información". Entonces, los términos de búsqueda deben ser precisos, porque si nuestro planteamiento es concreto, la consulta tendrá más sentido, estará más acotada y nos llevará a referencias apropiadas. Asimismo, nuestra búsqueda deberá hacerse con palabras en español y en inglés, porque gran cantidad de fuentes primarias se encuentran en este idioma.

Al consultar una base de datos, sólo nos interesan las referencias que se relacionen estrechamente con el problema específico que vamos a investigar. Por ejemplo, si pretendemos analizar la relación entre el clima organizacional y la satisfacción laboral, ¿cómo encontraremos las fuentes primarias que en verdad tienen que ver con el problema de estudio que nos incumbe? Primero, con la revisión de una base de datos apropiada. Si nuestro tema trata sobre clima organizacional y satisfacción laboral, *no* consultaríamos una base de referencias sobre temas de química, como *Chemical Abstracts,* ni una base de datos con referencias de la historia del arte, sino una base de información con fuentes primarias de la materia de estudio. Tal es el caso de Wiley (en *Business, Economics, Finance and Accounting*), *SAGE Journals*, *Communication Abstracts* y ABI/INFORM (bases de datos correctas para nuestra investigación). Si vamos a comparar diferentes métodos educativos por medio de un experimento, debemos acudir a la base de referencias adecuada: ERIC (Education Resources Information Center)[2]. En español también hay bases, como *Latindex* y *Redalyc*, para diversas ciencias y disciplinas; *bvs*, ciencias de la salud; *ENFISPO*, enfermería, etcétera).[3]

Una vez elegida la base de datos que emplearemos, procedemos a consultar el "catálogo de temas, conceptos y términos" (el tesauro) respectivo,[4] que contiene un diccionario o vocabulario en el cual podemos hallar un listado de palabras para realizar la búsqueda. Del catálogo debemos seleccionar las palabras o conceptos "clave" que le confieran una dirección a la consulta. También podemos hacer una *búsqueda avanzada* con esos términos, utilizando los operadores del *sistema booleano*: *and* (en español "*y*"), *or* (en español "*o*") y *not* (en español "*no*"). Con los descriptores y las preposiciones estableceremos los límites de la consulta al banco o la base de referencias.[5]

La búsqueda nos proporcionará un listado de referencias vinculadas a las palabras clave (dicho de otra manera, el listado que obtengamos dependerá de estos términos llamados descriptores, los cuales escogemos del diccionario o simplemente utilizamos los que están incluidos en el planteamiento). Por ejemplo, si nuestro interés se centra en los "procedimientos quirúrgicos para el cáncer de próstata en ancianos" y vamos a revisar en las bases de referencias "MEDLINE®/PubMed®/Resources Guide" (en

Centro de recursos en línea

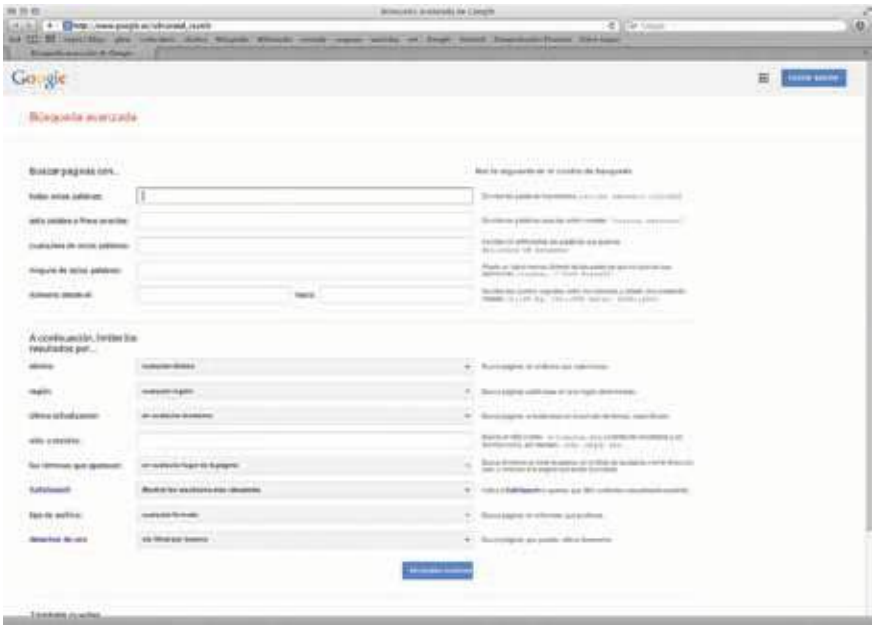
[2] Estas bases de referencias tienen páginas web en inglés y nuestra consulta requerirá básicamente de términos en este idioma.

[3] En el centro de recursos en línea Material complementario → Apéndices → Apéndice 2: "Principales bancos/servicios de obtención de fuentes/bases de datos/páginas web para consulta de referencias bibliográficas", el lector encontrará un listado de bases para sus búsquedas.

[4] De acuerdo con la Cornell University Library (2005), "el tesauro es una lista de todos los títulos o descriptores usados en cierta base de datos, catálogo o índice", proviene del latín thesaurus.

[5] Si usted, lector o lectora, no está familiarizado con estos operadores o preposiciones, o nunca ha hecho una "búsqueda avanzada", le sugerimos descargar y revisar el capítulo 3 del centro de recursos en línea (en la sección "Capítulos") titulado: "Perspectiva teórica: comentarios adicionales", donde se explican sus usos y funciones.

inglés) y "MedlinePlus®" (en español), si seleccionamos las palabras o descriptores "prostate cancer" y "cáncer de próstata", el resultado de la consulta será una lista de todas las referencias bibliográficas que estén en tales bases y que se relacionen con dichos términos (enfermedad). Si hicimos una búsqueda el 10 de abril de 2013 y obtuvimos 107 976 referencias en PubMed® y 492 en MedlinePlus® (que son demasiadas, por lo que tenemos que utilizar más descriptores o incrementar nuestra precisión). Al agregarle el término "old man" (consulta en PubMed) y "anciano" (búsqueda en MedlinePlus), el resultado fue de 2 055 y 24 referencias respectivamente (todavía muchas en la primera base). Y al agregarle "surgery" a la consulta en PubMed (porque realmente nuestro estudio se centra en ello), el número resultó mucho más manejable, 148 fuentes primarias (si en lugar de este descriptor usamos "surgical procedures", el número de referencias sería de 127). Desde luego, las búsquedas avanzadas pueden acotarse por fechas (por ejemplo, últimos tres años, de 2010 a 2016, de 2000 a 2015).



Google tiene uno de los mejores sistemas de búsquedas avanzadas, pero para una consulta más completa es necesario acudir a otras bases más especializadas, como EBSCO (Academic Search), SAGE, ERIC, Social Citation Index, Cochrane, JSTOR, PsycINFO, etc. (véase apéndice 2, del centro de recursos en línea).

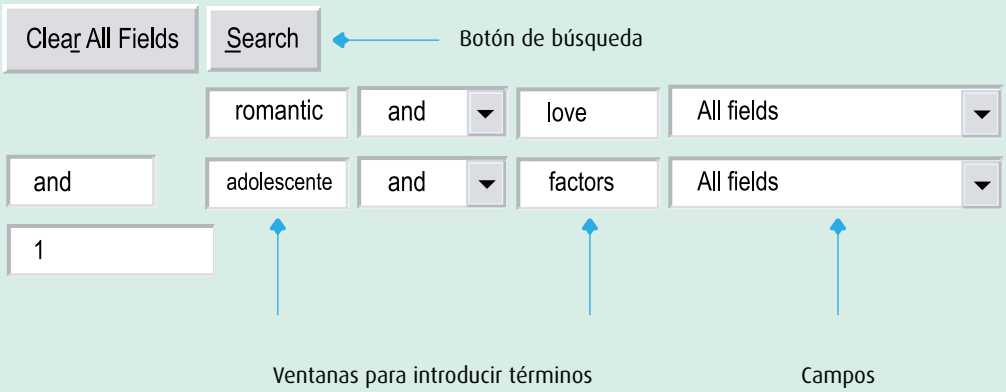
Ejemplo

En la mayoría de las bases de referencias o revistas existen dos opciones de búsqueda:

- Búsqueda sencilla ("search").
  Por lo general, aparece un recuadro o ventana donde se solicita introducir los términos de búsqueda, en este caso escribimos las palabras y damos clic donde corresponda para iniciar la búsqueda. Si entrecomillamos los descriptores sabemos que será literal como si usáramos el conector "and"(y).

- Búsqueda avanzada ("advanced search").
  En esta clase de búsqueda generalmente nos aparecen varias ventanas o recuadros para insertar los términos (uno por recuadro), además los *operadores booleanos* correspondientes y con frecuencia otra ventana para restringir la búsqueda por campo (autor, publicación, volumen, etc.; aunque la opción por *default* es: "todos los campos" —"all fields"—). Y en algún lugar se coloca el rango temporal de búsqueda (éste varía

en diferentes casos, es cuestión de ubicarlo y utilizarlo para restringir la consulta a un periodo: tal mes y año a tal mes y año —o simplemente de tal año a tal otro— o número de los últimos años).



Específicamente en lo que se refiere a libros, ya sabemos que podemos buscar en las páginas de las principales editoriales y librerías, así como en otros lugares (Amazon, AbeBooks en español, Books In Print®, etc.). De las referencias que encontremos en las búsquedas, elegimos las más convenientes (sobre esto se comentará más adelante).

Centro de recursos en línea

También hay *bancos de datos que se consultan manualmente*, en los que las referencias se buscan en libros. En el capítulo 3 del centro de recursos en línea (Material complementario → Capítulos → Capítulo "Perspectiva teórica: comentarios adicionales") se expone este tema y se presentan diversos ejemplos de búsquedas.

Consultar en internet es necesario y tiene ventajas, pero si *no* buscamos en sitios con verdadera información científica o académica de calidad, puede ser riesgoso. No es recomendable acudir a sitios con un fuerte uso comercial. Creswell (2013a) y Lester y Lester (2012) hacen un análisis de las ventajas y desventajas de utilizar internet en la búsqueda de literatura pertinente para el planteamiento del problema, como se ve en la tabla 4.1.

Amazon se ha posicionado como una de las librerías virtuales más completa en la red.

Tabla 4.1 Ventajas y desventajas de utilizar internet como fuente para localizar literatura

Ventajas	Desventajas
Acceso fácil las 24 horas del día	Con frecuencia las investigaciones colocadas en sitios web no están revisadas por expertos.
Gran cantidad de información en diversos sitios web sobre muchos temas	Los reportes de investigación u otros documentos de los sitios web pueden ser textos plagiados o que se reproducen sin el consentimiento de autores. Además del aspecto ético, al utilizarlos nuestro profesor, jefe, un comité, un cliente, etc., pueden pensar que todo nuestro trabajo es un plagio y reprobarnos, expulsarnos, demandarnos o suspendernos.
Información en español	Puede ser muy tardado localizar estudios de calidad sobre nuestro tema, pues abundan páginas o sitios que se refieren a nuestro planteamiento, pero no incluyen investigaciones con datos sino opiniones, ideas o servicios de consultoría.
Información reciente	La información puede estar desorganizada y ser poco útil.
El acceso a los sitios web es inmediato a través de buscadores	Para tener acceso a la mayoría de los textos completos de artículos, se debe pagar entre 5 y 30 dólares estadounidenses.
En la mayoría de los casos el acceso es gratuito o de muy bajo costo	
El investigador puede crear una red de contactos que le ayuden a obtener la información que busca	
Los estudios que se localicen pueden imprimirse de inmediato	

Palabras clave Para elegirlas se recomienda: escribir un título preliminar del estudio y seleccionar las dos o tres palabras que capten la idea central, extraer los términos del planteamiento o utilizar los que los autores más destacados en el campo de nuestro estudio utilicen en sus planteamientos e hipótesis. En la mayoría de los artículos de revistas es común incluir los términos clave al inicio o al final.

No todos los documentos de la web son dignos de tomarse en cuenta Debemos filtrar y no utilizar: 1) los que constituyen simples opiniones sin apoyo o resultan insustanciales, 2) aquellos que son comerciales y que incluyen información no corroborada y 3) otros que están sesgados y carecen de fundamentos.

Obtención (recuperación) de la literatura

Después de identificar las fuentes primarias pertinentes, es necesario localizarlas en las bibliotecas físicas y electrónicas, filmotecas, hemerotecas, videotecas u otros lugares donde se encuentren. Si compramos artículos de revistas científicas, los descargamos y guardamos en nuestro equipo para su posterior consulta (y también suelen imprimirse). Si son libros comprados en internet, estaremos pendientes de que lleguen a nuestras manos, etcétera.

Consulta de la literatura

Una vez que se han localizado físicamente las referencias (la literatura) de interés se procede a *consultarlas.* El primer paso consiste en seleccionar las que serán de utilidad para nuestro marco teórico específico y *desechar* las que *no* nos sirvan. En ocasiones, una fuente primaria puede referirse a nuestro problema de investigación, pero no sernos útil porque no enfoca el tema desde el punto de vista que pretendemos establecer, se han realizado nuevos estudios que han encontrado explicaciones más satisfactorias, que han invalidado sus resultados o que han desaprobado sus conclusiones, se detectaron errores de método o bien se realizaron en contextos completamente diferentes al de nuestra investigación, etc. Si la recopilación de la literatura se realiza mediante compilaciones o bancos de datos en los que se incluye un breve resumen de cada referencia, se corre menos riesgo de elegir una fuente primaria inútil.

En todas las áreas de conocimiento, las *fuentes primarias más consultadas y utilizadas* para elaborar marcos teóricos son *libros, artículos de revistas científicas y ponencias o trabajos presentados en congresos, simposios y eventos similares*, entre otras razones, porque estas fuentes son las que sistematizan en mayor medida la información, profundizan más en el tema que desarrollan y son altamente especializadas, además de que se puede tener acceso a ellas por internet. Así, Creswell (2013a) recomienda confiar en la medida de lo posible en artículos de revistas científicas, que son evaluados críticamente por editores y dictaminadores expertos antes de ser publicados. Tales artículos son

¿Qué buscamos al consultar la literatura? Básicamente: definiciones, teorías, resultados, casos, ejemplos, instrumentos utilizados para medir o evaluar los conceptos o variables de interés, hipótesis comprobadas, datos específicos y enfoques o abordajes al problema de investigación.

de los más diversos tipos y cada uno sirve para varios propósitos en el marco teórico, como puede apreciarse en la tabla 4.2. Su selección muchas veces depende de la clase de asignación que solicite a los alumnos el profesor, tutor o director de tesis.

Tabla 4.2 Principales tipos de artículos que podemos encontrar[6]

Tipos de artículos	Descripción del contenido básico	Utilidad potencial para nuestra investigación	Asignación del profesor para la cual es pertinente este tipo de artículos
Reportes de investigación empírica (resultados)	Presentan resultados y descubrimientos producto de recolectar y analizar datos, respondiendo a un planteamiento del problema.	Detectar o identificar palabras clave, planteamientos, abordajes o enfoques, teorías, hipótesis, variables, definiciones, diseños, muestras, instrumentos y análisis, así como conocer sus procedimientos. Encontrar ejemplos. Comparar nuestros resultados con estudios previos.	Trabajo o tesis que implica un estudio empírico (nos solicitan recolectar y analizar datos).
Reportes de evaluación o diagnóstico	Muestran los resultados de una evaluación o diagnóstico hecho a un programa, intervención, organización, proceso, maquinaria o el estado actual de uno o varios casos respecto de una o más variables.	Detectar o identificar palabras clave, planteamientos, abordajes o enfoques, teorías, hipótesis, variables, definiciones, diseños, muestras, instrumentos y análisis, así como conocer sus procedimientos (ejemplos: evaluar un proceso de calidad, un diagnóstico clínico, de la situación financiera de una empresa, del estado de una maquinaria). Encontrar ejemplos. Comparar nuestros resultados con estudios previos.	Trabajo o tesis que comprende una evaluación o diagnóstico (si es del estado actual, una sola recolección de datos; si se trata de evaluar una intervención, una recolección antes y otra después de ella).
Revisiones, síntesis, metaanálisis	Integran información de varias investigaciones previas o informes de evaluación. Consideran coincidencias y divergencias entre estudios previos para obtener conclusiones en el análisis de un planteamiento o fenómeno.	Detectar o identificar palabras clave, planteamientos, abordajes o enfoques, teorías, hipótesis, variables, definiciones, diseños, muestras, instrumentos y análisis, así como conocer sus procedimientos. Encontrar ejemplos. Comparar nuestros resultados con estudios previos.	Marco teórico de un trabajo o tesis que implica estudio empírico. Tesina o monografía. En el caso del metaanálisis: un estudio empírico amplio para analizar los datos obtenidos por investigaciones previas.
Artículos conceptuales o teoréticos	Analizan un nuevo concepto, variable, hipótesis o teoría; o bien, describen conceptos, variables, hipótesis o teorías existentes relacionadas con un planteamiento.	Detectar o identificar palabras clave, planteamientos, teorías, hipótesis, variables y definiciones.	Marco teórico de un trabajo o tesis que implica estudio empírico o teorético. Tesina o monografía.
Reportes descriptivos	Definen y describen un evento, fenómeno, hecho, programa, proceso, caso (individuo, objeto, organización, comunidad, etcétera).	Detectar o identificar palabras clave, planteamientos, conceptos, variables y definiciones, muestras, instrumentos y análisis. Encontrar ejemplos. Comparar nuestros resultados con estudios previos.	Trabajo o tesis que implica un estudio empírico descriptivo (nos piden recolectar y analizar datos).
Ensayos	Presentan una argumentación respaldada por evidencias.	Detectar o identificar palabras clave, planteamientos, conceptos, variables y definiciones potenciales. Encontrar ejemplos. Interpretar nuestras conclusiones.	Monografía. Elaborar la introducción de un trabajo o tesis que implica un estudio empírico o teorético. Desarrollo de la discusión (conclusiones de un trabajo o tesis).
De opiniones	Exponen el punto de vista de un autor respecto a un tópico.	Detectar o identificar palabras clave, planteamientos, conceptos y variables potenciales.	Introducción de un trabajo o tesis.
Revisiones de libros o artículos	Describen y analizan los elementos básicos de libros y artículos (objetivos, contenidos básicos, enfoque y utilidad).	Detectar libros y artículos que pueden ser útiles para nuestra propia investigación.	Regularmente no se integran a las referencias o bibliografía, pues simplemente ayudan en la decisión sobre consultar el libro original.

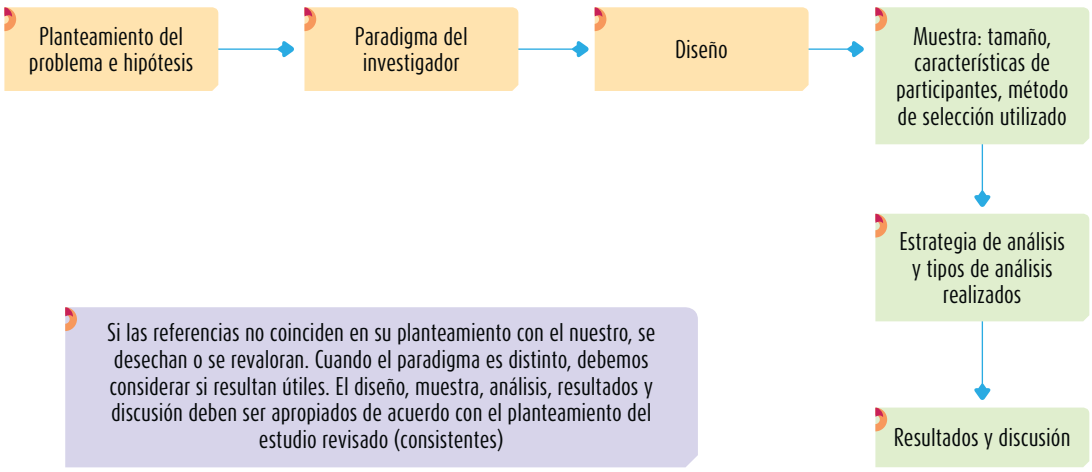
[6] Adaptado parcialmente de Savin-Baden y Major (2013, p. 118).

www.elosopanda.com | jamespoetrodriguez.com

Para establecer la utilidad de los libros, conviene comenzar analizando el índice de contenido y el índice analítico o de materias por cuestión de tiempo, los cuales proporcionan una idea de los temas incluidos en la obra. Al tratarse de artículos de revistas científicas, lo más adecuado es revisar primero el resumen y palabras clave, y en caso de considerarlo de utilidad, examinar las conclusiones, observaciones o comentarios finales o, en última instancia, todo el artículo.

Mertens (2010) y Creswell (2013a) proponen una revisión que se aplica a prácticamente cualquier referencia, que implica recolección y análisis de los datos y se presenta en la figura 4.1.

Figura 4.1 Revisión de una referencia primaria.



Con el propósito de seleccionar las fuentes primarias que servirán para elaborar el marco teórico, es conveniente hacerse las siguientes preguntas:

- ¿La referencia se relaciona con mi problema de investigación? ¿Cómo?
- ¿Qué aspectos trata?
- ¿Ayuda a que se realice más rápido y profundamente mi estudio?
- ¿Desde qué óptica y perspectiva aborda el tema (psicológica, antropológica, sociológica, médica, legal, económica, comunicológica, administrativa, de ingeniería industrial, etcétera)?

La respuesta a esta última pregunta es muy importante. Por ejemplo, si se pretende estudiar la relación entre superior y subordinado en términos del efecto que la realimentación positiva del primero tiene en la motivación del segundo, la investigación posee un enfoque principalmente comunicológico. Supongamos que encontramos un artículo que versa sobre la relación entre un superior o jefe y su subordinado, pero trata de las atribuciones administrativas que cierto tipo de subordinados tienen en determinadas empresas. Resulta obvio que este artículo se debe descartar, pues enfoca el tema desde otra perspectiva.

Lo anterior no significa que no se pueda acudir a otros campos de conocimiento para completar la revisión de la literatura, pues en algunos casos se encuentran referencias sumamente útiles en otras áreas.

Para analizar las referencias, recordemos que se toma en cuenta:

- Cercanía o similitud a nuestro planteamiento (utilidad).
- Semejanza a nuestro método y muestra.
- Fecha de publicación o difusión (cuanto más reciente, mejor).
- Que consista en una investigación empírica (recolección y análisis de datos).
- Rigor y calidad del estudio (cuantitativo, cualitativo o mixto).

Por lo que se refiere al *apoyo bibliográfico*, algunos investigadores consideran que no debe acudirse a obras preparadas en el extranjero, porque la información que presentan y las teorías que sostienen

fueron elaboradas para otros contextos y situaciones. Aunque esto es cierto, no implica que deba rechazarse o no utilizarse tal material; la cuestión es saber cómo usarlo. Es posible que la *literatura extranjera* le ayude al investigador local de diversas formas: puede ofrecerle un buen punto de partida, guiarlo en el enfoque y tratamiento que se le dará al problema de investigación, orientarlo respecto de los diversos elementos que intervienen en el problema, centrarlo en un problema específico, sugerirle cómo construir el marco teórico, etcétera.

Un caso ilustrativo fueron los estudios de Rota (1978), cuyo propósito primordial era analizar el efecto que la exposición a la violencia televisada tiene en la conducta agresiva de los niños. Cuando el autor citado revisó la literatura, encontró que prácticamente no se habían realizado estudios previos en México, pero que en Estados Unidos se habían llevado a cabo investigaciones y que incluso se habían formulado teorías al respecto (teoría del reforzamiento, teoría de la catarsis y las teorías de los efectos disfuncionales). El autor se basó en la literatura estadounidense y comenzó a efectuar estudios en México. Sus resultados variaron de los encontrados en Estados Unidos, aunque los antecedentes localizados en esa nación constituyeron un excelente marco de referencia y un punto de partida para sus investigaciones.

Desde luego, en ocasiones ciertos fenómenos evolucionan o cambian al paso del tiempo. Por ejemplo, podría ser que una generación de niños no se viera influida por ciertos efectos de la televisión y otra generación sí, lo cual quiere decir que las ciencias no son estáticas. Hoy en día, nuestras percepciones y conocimientos sobre diversos fenómenos han cambiado con el desciframiento del genoma humano, los actos terroristas de 2001 en Estados Unidos, el huracán Sandy que impactó en 2012 Nueva York y diversas ciudades aledañas (también Haití, Cuba y otros países del Caribe), las expediciones más recientes al espacio, las últimas guerras, el desarrollo de las comunicaciones telefónicas o los sucesos locales.

Una vez seleccionadas las referencias o fuentes primarias útiles para el problema de investigación, se revisan cuidadosamente y se extrae la información necesaria para integrarla y desarrollar el marco teórico. Al respecto, es recomendable anotar los datos completos de identificación de la referencia.[7]

¿Qué información o contenido se extrae de las referencias?

Centro de recursos en línea

A veces se extrae una sola o varias ideas; otras, una cifra, un resultado o numerosos comentarios (algunos ejemplos y su organización se muestran en el centro de recursos en línea de la obra, en el capítulo 3, "Perspectiva teórica: comentarios adicionales").[8]

Al identificar la literatura útil, se puede diseñar un *mapa de revisión*, el cual es una "imagen de conceptos" de la agrupación propuesta respecto a las referencias del planteamiento y que ilustra cómo la indagación contribuirá al estudio (veremos un ejemplo más adelante).

Cuando ya se haya reunido la literatura que se consideró para la elaboración del *mapa de revisión*, también se deben empezar a generar los resúmenes de los artículos y documentos más relevantes y la extracción de ideas, cifras y comentarios. Estos resúmenes e información se combinarán posteriormente en el marco teórico (Hernández-Sampieri y Méndez, 2009).

¿Qué nos puede revelar la revisión de la literatura?

Uno de los propósitos de la revisión de la literatura es analizar y discernir si la teoría y la investigación anterior sugieren una respuesta (aunque sea parcial) a la pregunta o las preguntas de investigación, o bien si provee una dirección a seguir dentro del planteamiento de nuestro estudio (Lawrence, Machi y McEvoy, 2012; Race, 2008).

Centro de recursos en línea

[7] En el centro de recursos en línea de la obra (Manuales auxiliares / Manual APA abreviado), el lector encontrará un pequeño manual basado en las normas de la APA (American Psychological Association) que se usa en la mayoría de las disciplinas. Ahí se señala qué elementos se anotan de las principales referencias y cómo citarlas en la lista final de referencias o bibliografía. Además, en dicha página podrá bajar el programa SISI (Sistema de Información para el Soporte a la Investigación) y su respectivo manual (en Manuales Auxiliares), herramientas que sirven para generar, incluir y organizar referencias bibliográficas, tanto en el texto —citas— como al final en el listado o bibliografía —referencias—, basándonos en el estilo de publicación de la APA.

[8] Para esta sexta edición se agregó un esquema de extracción de información basado en tablas o matrices que puede resultarle útil al lector.

La literatura revisada puede revelar diferentes grados en el desarrollo del conocimiento:

- Que hay una teoría completamente desarrollada, con abundante evidencia empírica[9] y que se aplica a nuestro problema de investigación.
- Que hay varias teorías con soporte empírico que se aplican a nuestro problema de investigación.
- Que hay "piezas y trozos" de teoría con cierto respaldo empírico, que sugieren variables potencialmente importantes y que se aplican a nuestro problema de investigación (pueden ser generalizaciones empíricas e hipótesis con apoyo de algunos estudios).
- Que hay descubrimientos interesantes, pero parciales, sin llegar a ajustarse a una teoría.
- Que sólo hay guías aún sin estudiar e ideas vagamente relacionadas con el problema de investigación.

Asimismo, nos podemos encontrar que los estudios antecedentes presentan falta de congruencia o claridad, debilidades en el método (en sus diseños, muestras, instrumentos para recolectar datos, etc.), aplicaciones que no han podido implementarse correctamente o que han mostrado problemas (Grobbee y Hoes, 2014; Mertens, 2010).

En cada caso, varía la estrategia que habremos de utilizar para *construir* y *organizar nuestro marco teórico*.

1. *Existencia de una teoría completamente desarrollada*

Cuando la revisión de la literatura revela que hay una teoría capaz de describir, explicar y predecir el planteamiento o fenómeno de estudio de manera lógica, completa, profunda y coherente, la mejor estrategia para construir el marco teórico es tomar esa teoría como la estructura misma.

En términos generales, una **teoría** es un conjunto de proposiciones interrelacionadas capaces de explicar por qué y cómo ocurre un fenómeno. En palabras de Kerlinger y Lee (2002), la teoría constituye un conjunto de constructos (conceptos) vinculados, definiciones y proposiciones que presentan una visión sistemática de los fenómenos al especificar las relaciones entre variables, con el propósito de explicar y predecir los fenómenos.

Teoría Conjunto de proposiciones interrelacionadas capaces de explicar por qué y cómo ocurre un fenómeno.

Las teorías pueden estar más o menos desarrolladas y tener mayor o menor valor. Los criterios para evaluarlas, así como una explicación e ilustración de éstas y las concepciones erróneas respecto de lo que es una teoría, las podrá encontrar el lector en el centro de recursos en línea de esta obra, capítulo 3, "Perspectiva teórica: comentarios adicionales".

Centro de recursos en línea

Ahora bien, si se descubre una teoría que explica el problema de investigación que nos interesa, se debe tener cuidado de no investigar algo ya estudiado muy a fondo. Imaginemos que alguien pretende realizar una investigación para someter a prueba la siguiente hipótesis referente al sistema solar: "Las fuerzas centrípetas tienden a los centros de cada planeta" (Newton, 1984, p. 61). Sería ridículo, porque es una hipótesis generada hace más de 300 años, comprobada de modo exhaustivo y que ha pasado a formar parte del saber común.

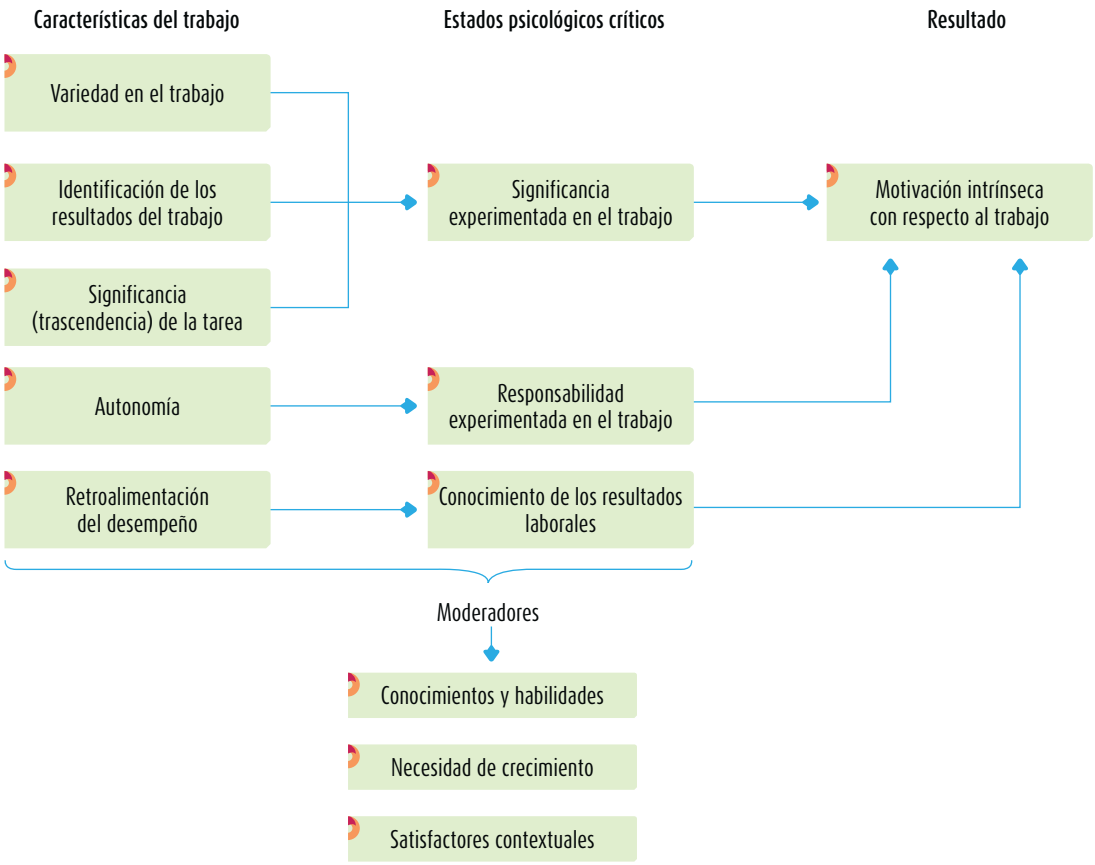
Cuando encontramos una teoría sólida que explica el planteamiento de interés, debemos darle un nuevo enfoque a nuestro estudio: a partir de lo que ya está comprobado, plantear otras interrogantes de investigación, obviamente aquellas que no ha podido resolver la teoría; o bien nos ocupamos de profundizar y ampliar elementos de la teoría y visualizar nuevos horizontes. También puede haber una buena teoría, pero aún no está comprobada o no se ha aplicado a nuestro contexto. De ser así, resultaría de interés someterla a prueba empírica en otras condiciones. Por ejemplo, una teoría de las causas de la satisfacción laboral formulada en Japón, que deseamos probar en Argentina o Brasil; o una teoría de los efectos de la exposición a contenidos sexuales en la televisión que únicamente se haya investigado en adultos, pero no en adolescentes.

En el caso de una teoría desarrollada, nuestro marco teórico consistirá en explicar la teoría, ya sea proposición por proposición o en forma cronológica para mostrar su evolución. Supongamos que se

[9] La evidencia empírica, en el enfoque cuantitativo, se refiere a los datos de la "realidad" que apoyan o dan testimonio de una o varias afirmaciones. Se dice que una teoría ha recibido apoyo o evidencia empírica cuando hay investigaciones científicas que han demostrado que sus postulados son ciertos en la realidad observable o medible. Las proposiciones o afirmaciones de una teoría llegan a tener diversos grados de evidencia empírica: *a*) si no hay evidencia empírica a favor ni en contra de una afirmación, a ésta se le denomina "hipótesis"; *b*) si hay apoyo empírico, pero éste es moderado, la afirmación o proposición es una "generalización empírica", y *c*) si la evidencia empírica es abrumadora, hablamos de "ley" (Sullivan, 2009; Reynolds, 1980).

intenta responder la siguiente pregunta: ¿cuáles son las características del trabajo relacionadas con la motivación para terminar las tareas laborales?[10] Al revisar la literatura, se encontraría una teoría sumamente desarrollada, designada como la teoría de la relación entre las características del trabajo y la motivación intrínseca. Esta teoría puede resumirse en el modelo de la figura 4.2 (adaptado de Hackman y Oldham, 1980, p. 83).[11]

Figura 4.2 Moderadores de la relación entre las características de trabajo y la motivación intrínseca.



Nuestro marco teórico se basaría en dicha teoría, incorporándole ciertas referencias de interés. Algunos autores lo estructurarían de la siguiente manera:

1. La motivación intrínseca con respecto al trabajo.
   - 1.1 ¿Qué es la motivación intrínseca en el contexto laboral?
   - 1.2 La importancia de la motivación intrínseca en el trabajo: su relación con la productividad.
2. Los factores del trabajo.
   - 2.1 Factores organizacionales (clima organizacional, políticas de la empresa, instalaciones, características estructurales de la organización: tamaño, tecnología, normas de la organización, entre otras cuestiones). *[Tratados de forma muy breve porque la investigación se enfoca en otros aspectos.]*
   - 2.2 Factores del desempeño (atribuciones internas, sentimientos de competencia y autodeterminación, etc.) *[También tratados muy brevemente por la misma razón.]*

[10] En un contexto determinado, digamos empresas del Parque Industrial de Villa El Salvador, en Lima, Perú.
[11] Este modelo sigue siendo utilizado, véase, por ejemplo: Hernández-Sampieri (2005), Østhus (2007), Prowse y Prowse (2008), Russell (2008), Rothausen (2009), Harley, Sargent y Allen (2010); Dierdorff, Rubin y Bachrach (2012) y Steger *et al.* (2013).

www.elosopanda.com | jamespoetrodriguez.com

- 2.3 Factores personales (conocimientos y habilidades, interés inicial por el trabajo y variables de personalidad, necesidades de desarrollo, etc.). *[Igualmente tratados en forma abreviada.]*
- 2.4 Factores de recompensa extrínseca (salario, prestaciones y otros tipos de recompensas). *[Comentados muy concisamente.]*

3. Características del trabajo.
   - 3.1 Variedad en el trabajo.
   - 3.2 Identificación de los resultados del individuo en el producto final.
   - 3.3 Importancia o trascendencia del trabajo.
   - 3.4 Autonomía.
   - 3.5 Realimentación del desempeño.
     - 3.5.1 Realimentación proveniente de agentes externos (superiores, supervisión técnica y compañeros de trabajo, que también constituyen una forma de recompensa extrínseca).
     - 3.5.2 Realimentación proveniente del trabajo en sí.
   - 3.6 Otras características.
4. La relación entre las características del trabajo y la motivación intrínseca. *[Aquí se comentaría cómo se vinculan entre sí dichas características y la forma en que se asocian, como un todo, a la motivación intrínseca. En esta parte del marco teórico, las características del trabajo se tomarían en conjunto, mientras que en el inciso 3 se menciona su correlación individual con la motivación intrínseca. Es decir, se explicaría el modelo de los moderadores de la relación entre las características del trabajo y la motivación intrínseca, a manera de resumen.]*

En este caso, por lo menos alrededor de 80% del marco teórico se desarrollaría en los incisos 3 y 4. De hecho, el inciso 2 es narrativo y general, y podría eliminarse. Su papel se limita a centrar el estudio en las variables de interés. En lo personal, *se agruparían* los factores organizacionales, del desempeño, personales y de recompensa extrínseca en un solo apartado, puesto que de ellos sólo se hablará en términos muy generales. Así, obtendríamos un capitulado más sencillo.

Otra perspectiva de nuestro marco teórico es la *cronológica*, que consiste en desarrollar históricamente la evolución de la teoría (analizar las contribuciones más importantes al problema de investigación hasta llegar a la teoría resultante). Si lo desarrolláramos siguiendo una perspectiva cronológica, tendríamos una estructura como la siguiente:

1. La motivación intrínseca y la motivación extrínseca: una división de la motivación hacia el trabajo.
2. Los modelos motivacionales clásicos para estudiar la motivación intrínseca.
   - 2.1. Antecedentes — Década de 1950
   - 2.2. Frederick Herzberg — Década de 1960
   - 2.3. Victor Vromm — De la década de 1950 a la de 1970
   - 2.4. Edward E. Lawler — De la década de 1960 a la de 1970
   - 2.5. Edward L. Deci — De la década de 1970 a la de 1990
3. El modelo de rediseño del trabajo (Richard Hackman y Greg Oldham)

   De la década de 1980 a la fecha
4. Las nuevas redefiniciones: Richard Ryan y Edward Deci, Kenneth W. Thomas, Steven Reiss.[12]

   De 2000 a la fecha

En los apartados se hablaría de las características del trabajo consideradas por cada autor o perspectiva en particular, así como su relación con la motivación intrínseca. *[Aunque el punto dos sería tratado muy brevemente.]* Al final se incluiría la teoría resultante, producto de años de investigación. Ya sea que decidamos construir el marco teórico cronológicamente o desglosar la estructura de la

[12] Hon (2012), Reiss (2012) y Chen (2013).

teoría (tratando, una por una, las proposiciones y los elementos principales de ella). Lo importante es explicar con claridad la teoría y la forma como se aplica a nuestro problema de investigación.

1. ***Existencia de varias teorías aplicables a nuestro problema de investigación***

Cuando al revisar la literatura se descubren varias teorías o modelos aplicables al problema de investigación, podemos elegir uno y basarnos en éste para edificar el marco teórico (desglosando la teoría o de manera cronológica); o bien, tomar partes de algunas o todas las teorías.

En la primera situación, elegimos la teoría que reciba una evaluación más positiva (de acuerdo con los criterios para evaluar una teoría que se comentan en el capítulo 3 del centro de recursos en línea)[13] y que se aplique más al problema de investigación. Por ejemplo, si el planteamiento se centra en los efectos que tienen en los adolescentes los programas televisivos con alto contenido sexual, podríamos encontrar teorías que expliquen el efecto de presenciar escenas de sexo en televisión, pero sólo una tiene que ver con adolescentes o cuenta con evidencia empírica del contexto elegido. Sin duda, ésta debería ser la teoría que seleccionemos para construir nuestro marco teórico.

En la segunda situación, se tomaría de las teorías sólo aquello que se relaciona con el problema de estudio. En estos casos, antes de trazar el marco teórico conviene bosquejarlo, analizarlo y decidir qué se va a incluir de cada teoría, procurando no caer en contradicciones lógicas (en ocasiones diversas teorías rivalizan en uno o más aspectos de manera total; si aceptamos lo que dice una teoría debemos desechar lo que postulan las demás). Cuando las proposiciones más importantes de las teorías se excluyen unas a otras, se debe elegir una sola. Pero si únicamente difieren en aspectos secundarios, se toman las proposiciones centrales que son más o menos comunes a todas ellas, y se eligen las partes de cada teoría que sean de interés y se acoplen entre sí.

Lo más común para construir el marco teórico es tomar una teoría como base y extraer elementos de otras teorías útiles.[14]

2. ***Existencia de "piezas y trozos" de teorías (generalizaciones empíricas)***

En ciertos campos del conocimiento no se dispone de muchas teorías que expliquen los fenómenos que estudian; a veces sólo se tienen **generalizaciones empíricas**, es decir, proposiciones que han sido comprobadas en la mayor parte de las investigaciones realizadas. Al revisar la literatura, es muy probable encontrar una situación así. Lo que se hace entonces es construir la perspectiva teórica, más que adoptar o adaptar una o varias teorías.

Generalizaciones empíricas Proposiciones que han sido comprobadas en la mayor parte de las investigaciones realizadas (constituyen la base de lo que serán las hipótesis que se someterán a prueba).

Cuando al revisar la literatura se encuentra una proposición única o en el planteamiento se piensa limitar la investigación a una generalización empírica (hipótesis), el marco teórico abarca los resultados y las conclusiones a que han llegado los estudios antecedentes, de acuerdo con algún esquema lógico (de manera cronológica, por variable o concepto de la proposición, o por las implicaciones de las investigaciones anteriores). Pero recordemos que nuestro estudio debe innovar.[15] Si nuestra pregunta de investigación fuera: ¿los individuos de un sistema social que adoptan primero una innovación están más expuestos a los canales interpersonales de comunicación que quienes la adoptan con posterioridad?,[16] nuestro marco teórico consistiría en comentar los estudios de difusión de innovaciones que, de una u otra manera, han hecho referencia al problema de investigación. Comentar implicaría describir cada estudio, el contexto en que se realizó y los resultados y las conclusiones a las que se llegó.

Ahora bien, casi todos los estudios se plantean varias preguntas de investigación o una pregunta de la cual se derivan diversas proposiciones (desde luego, como se mencionó en el capítulo anterior, evitando una dispersión incoherente). En estos casos, el marco teórico también se fundamentaría en

[13] Estos criterios son: *a*) capacidad de descripción, explicación y predicción , *b*) congruencia lógica, *c*) perspectiva, *d*) innovación e inventiva (fructificación heurística) y *e*) sencillez (parsimonia).

[14] Para ver cómo se integra un marco teórico en torno a una teoría o un modelo, sugerimos al lector descargar del centro de recursos en línea y revisar el ejemplo 6, "Exploración factorial del clima y la cultura organizacional en el marco del modelo de los valores en competencia" (en "Material complementario", "Investigación cuantitativa").

[15] A veces se llevan a cabo investigaciones para evaluar la falta de coherencia entre estudios previos, encontrar "huecos" de conocimiento en éstos o explorar por qué ciertas aplicaciones no han podido implantarse adecuadamente.

[16] Tomada de Rogers y Shoemaker (1971). Ejemplos de innovaciones son la moda, la tecnología, los sistemas de trabajo, etc.; y desde luego, debe especificarse el sistema social.

los estudios anteriores que se refieren a tales proposiciones. Los estudios se comentan y se van relacionando unos con otros, de acuerdo con un criterio coherente (cronológicamente, por cada proposición o por las variables del estudio). En ocasiones se entrelazan las proposiciones de manera lógica para postular, en forma tentativa, una teoría (la investigación puede comenzar a integrar una teoría que se encargarán de afinar estudios futuros).

Cuando nos encontramos con generalizaciones empíricas, es frecuente organizar el marco teórico por cada una de las variables del estudio. Por ejemplo, si pretendemos investigar el efecto que producen ciertas dimensiones del clima organizacional sobre la rotación de personal, nuestro marco teórico podría tener la siguiente estructura:

1. Definiciones fundamentales: el clima organizacional y la rotación de personal.
2. Dimensiones del clima organizacional[17] y su efecto en la rotación de personal.
   - 2.1 Moral
   - 2.2 Apoyo de la dirección
   - 2.3 Motivación intrínseca
   - 2.4 Autonomía
   - 2.5 Identificación con la organización
   - 2.6 Satisfacción laboral

En cada subsección del inciso 2 se definiría la dimensión y se incluirían las generalizaciones o proposiciones empíricas sobre la relación entre la variable y la rotación.

Las generalizaciones empíricas que se descubran en la literatura constituyen la base de lo que serán las hipótesis que se someterán a prueba y a veces son las hipótesis mismas. Lo mismo ocurre cuando tales proposiciones forman parte de una teoría.

3. ***Descubrimientos interesantes pero parciales que no se ajustan a una teoría***

En la literatura podemos encontrar que no hay teorías ni generalizaciones empíricas, sino sólo algunos estudios previos vinculados —relativamente— con nuestro planteamiento. Podemos organizarlos como antecedentes de forma lógica y coherente, destacando lo más relevante en cada caso y citándolos como puntos de referencia. Se debe ahondar en lo que cada antecedente aporta.

Por ejemplo, el estudio ya mencionado en el capítulo anterior, de Núñez (2001), quien finalmente diseñó una investigación para validar un instrumento que midiera el sentido de vida de acuerdo con el pensamiento y la filosofía de Viktor Frankl. Al revisar la literatura se encontró con que había otras pruebas logoterapéuticas que medían el propósito de vida; pero que no reflejaban totalmente el pensamiento de dicho autor. Construyó su marco teórico alrededor del modelo concebido por Frankl (manifestaciones del espíritu, libertad, responsabilidad, conciencia, valores, etc.) y tomó los instrumentos previos como puntos de referencia. No se *acogió* a una teoría, *adaptó* un esquema de pensamiento y enmarcó su estudio con otros anteriores (que desarrollaron diversos instrumentos de medición). Entre algunos de sus apartados del marco teórico incluyó puntos como los siguientes:

Medición del sentido de vida:

- Tests logoterapéuticos
- El test de propósito vital de Crumbaugh y Maholick (PIL)
- Investigaciones realizadas con el PIL
- Investigaciones en México
- Test de Song
- Escala de vacío existencial (EVS) del MMPI
- Cuestionario de propósito vital (LPQ)
- El test del significado del sufrimiento de Starck
- Test de Belfast
- Logo-test de Elizabeth Lukas

[17] Se simplificaron las dimensiones del clima organizacional para hacer más ágil el ejemplo.

4. ***Existencia de guías aún no investigadas e ideas vagamente relacionadas con el problema de investigación***

En ocasiones se descubre que se han efectuado pocos estudios dentro del campo de conocimiento en cuestión. En tales casos, el investigador tiene que buscar literatura que, aunque no se refiera al problema específico de la investigación, lo ayude a orientarse dentro de él. Paniagua (1985), al llevar a cabo una revisión de la bibliografía sobre las relaciones interpersonales del comprador y el vendedor en el contexto organizacional mexicano, no detectó ninguna fuente primaria sobre el tema específico. Entonces, tomó referencias de relaciones interpersonales provenientes de otros contextos (superior y subordinado, entre compañeros de trabajo y desarrollo de las relaciones en general), y las aplicó a la relación entre el comprador y el vendedor para construir el marco teórico.

Tomemos otro caso para ilustrar cómo se constituye el marco teórico en situaciones en las que no hay estudios previos sobre el problema de investigación específico. Supongamos que se trata de analizar qué factores del contexto laboral provocan el temor a tener logros[18] y repercuten en la motivación de logro de las secretarias que trabajan en la burocracia gubernamental de Costa Rica. Quizá se encuentre que no hay ningún estudio al respecto, pero tal vez existan investigaciones sobre el temor y la motivación de logro de las secretarias costarricenses (aunque no laboren en el gobierno) o de supervisores de departamentos públicos (aunque no se trate de la ocupación que específicamente nos interesa). Si tampoco ocurre lo segundo, tal vez haya estudios que tratan ambas variables con ejecutivos de empresas privadas o de secretarias de dependencias públicas de otros países. Si no es así, se acude a las investigaciones sobre el temor y la motivación de logro, a pesar de que probablemente se hayan realizado entre estudiantes de otro país. Pero, si no hubiera ningún antecedente se recurriría a los estudios iniciales de motivación de logro de David McClelland y a los de temor del logro. Aunque, por ejemplo, para temor de logro encontraríamos múltiples referencias (Nguyen y Benet-Martínez, 2013; Roth, Purvis y Bobko, 2012; Choi, 2011; Gore *et al.,* 2011; Romer, 2010, Mulig *et al.,* 2006; Kocovski y Endler, 2000; Lew *et al.,* 1998; Janda, O'Grady y Capps, 1978; Cherry y Deaux, 1978, Tresemer, 1977 y 1976, y Zuckerman *et al.,* 1980; entre otras). Pero en el supuesto de que tampoco las hubiera, se acudiría a estudios generales sobre temor y motivación. Sin embargo, casi siempre se cuenta con un punto de partida. Las excepciones en este sentido son muy pocas. Las quejas de que "no hay nada", "nadie lo ha estudiado", "no sé en qué antecedentes puedo basarme", por lo general se deben a una deficiente revisión de la literatura. Se encuentra otro ejemplo sobre qué hacer cuando no hay literatura (incluso sobre cuestiones no inventadas) en el ya referido capítulo 3 que se puede descargar del centro de recursos en línea.

Algunas observaciones sobre el desarrollo de la perspectiva teórica

En el proceso cuantitativo siempre es conveniente efectuar la revisión de la literatura y presentarla de una manera organizada (llámese marco teórico, marco de referencia, conocimiento disponible o de cualquier otro modo), y aunque nuestra investigación puede centrarse en un objetivo de evaluación o medición muy específico (por ejemplo, un estudio que solamente pretenda medir variables particulares, como el caso de un censo demográfico en una determinada comunidad donde se medirían: nivel socioeconómico, nivel educativo, edad, género, tamaño de la familia, etc.), es recomendable revisar lo que se ha hecho antes (cómo se han realizado en esa comunidad los censos demográficos anteriores o, si no hay antecedentes en ella, cómo se han efectuado en comunidades similares; qué problemas se tuvieron, cómo se resolvieron, qué información relevante fue excluida, etc.). Esto ayudará a concebir un estudio mejor y más completo. En investigaciones para mejorar procesos, desarrollar algoritmos y generar nuevas tecnologías y productos, parte de la revisión de la literatura consiste en considerar y evaluar los anteriores procesos, algoritmos, tecnologías o productos, recordando que el planteamiento implica analizar en qué variables puede optimizarse.

[18] Temor a ser exitoso en un trabajo u otra tarea.

www.elosopanda.com | jamespoetrodriguez.com

Rango Extensión total de los datos en la escala.

El **rango**, también llamado *recorrido*, es la diferencia entre la puntuación mayor y la puntuación menor, e indica el número de unidades en la escala de medición que se necesitan para incluir los valores máximo y mínimo. Se calcula así: $X_M - X_m$ (puntuación mayor menos puntuación menor). Si tenemos los siguientes valores:

17 18 20 20 24 28 28 30 33

El rango será: 33 – 17 = 16.

Cuanto *más grande* sea el *rango*, *mayor* será la *dispersión de los datos* de una distribución.

Desviación estándar Promedio de desviación de las puntuaciones con respecto a la media que se expresa en las unidades originales de medición de la distribución.

La **desviación estándar** o característica es el promedio de desviación de las puntuaciones con respecto a la media (Jarman, 2013 y Levin, 2003). Esta medida se expresa en las unidades originales de medición de la distribución. Se interpreta en relación con la media. Cuanto mayor sea la dispersión de los datos alrededor de la media, mayor será la desviación estándar. Se simboliza como: s o la sigma minúscula σ, o bien mediante la abreviatura DE. Su cálculo lo podrá encontrar el lector en el centro de recursos, en: Material complementario → Documentos → Documento 2, "Fórmulas y procedimientos estadísticos".

Centro de recursos en línea

La desviación estándar se interpreta como *cuánto se desvía, en promedio, de la media un conjunto de puntuaciones*.

Supongamos que un investigador obtuvo para su muestra una media (promedio) de ingreso familiar anual de 6 000 unidades monetarias y una desviación estándar de 1 000. La interpretación es que los ingresos familiares de la muestra se desvían, en promedio, mil unidades monetarias respecto a la media.

La desviación estándar sólo se utiliza en variables medidas por intervalos o de razón.

La varianza

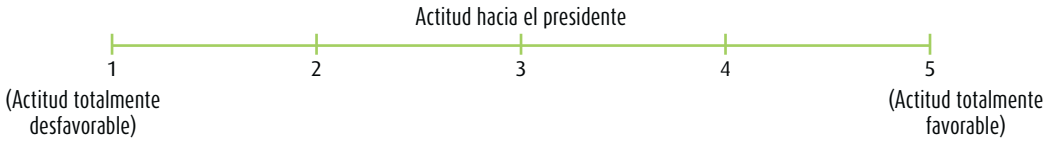
Varianza Desviación estándar elevada al cuadrado.

La **varianza** es la desviación estándar elevada al cuadrado y se simboliza como s^2. Es un concepto estadístico muy importante, ya que la mayoría de las pruebas cuantitativas se fundamentan en él. Diversos métodos estadísticos parten de la descomposición de la varianza (Zhang, 2013; Beins y McCarthy, 2009; Wilcox, 2008; y Jackson, 2008). Sin embargo, con fines descriptivos se utiliza preferentemente la desviación estándar.

¿Cómo se interpretan las medidas de tendencia central y de la variabilidad?

Cabe destacar que al describir nuestros datos, respecto a cada *variable del estudio*, interpretamos las medidas de tendencia central y de la variabilidad en conjunto, no aisladamente. Consideramos todos los valores. Para interpretarlos, lo primero que hacemos es tomar en cuenta el rango potencial de la escala. Supongamos que aplicamos una escala de actitudes del tipo Likert para medir la "actitud hacia el presidente" de una nación (digamos que la escala tuviera 18 ítems y se promediaran sus valores). El rango potencial es de uno a cinco (véase la figura 10.7).

Figura 10.7 Ejemplo de escala con rango potencial.



Si obtuviéramos los siguientes resultados:

Variable: actitud hacia el presidente
Moda: 4.0

Mediana: 3.9
Media ($\overline{X}$): 4.2
Desviación estándar: 0.7
Puntuación más alta observada (máximo): 5.0
Puntuación más baja observada (mínimo): 2.0
Rango: 3

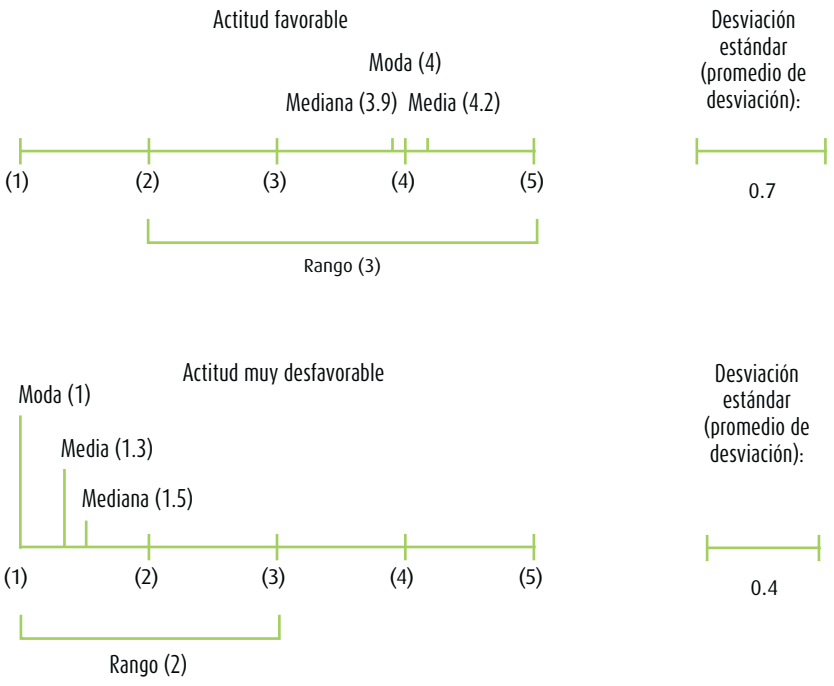
podríamos hacer la siguiente interpretación descriptiva: la actitud hacia el presidente es favorable. La categoría que más se repitió fue 4 (favorable). Cincuenta por ciento de los individuos está por encima del valor 3.9 y el restante 50% se sitúa por debajo de este valor (mediana). En promedio, los participantes se ubican en 4.2 (favorable). Asimismo, se desvían de 4.2, en promedio, 0.7 unidades de la escala. Ninguna persona calificó al presidente de manera muy desfavorable (no hay "1"). Las puntuaciones tienden a ubicarse en valores medios o elevados.

En cambio, si los resultados fueran:

Variable: actitud hacia el presidente
Moda: 1
Mediana: 1.5
Media ($\overline{X}$): 1.3
Desviación estándar: 0.4
Máximo: 3.0
Mínimo: 1.0
Rango: 2.0

La interpretación es que la actitud hacia el presidente es muy desfavorable. En la figura 10.8 vemos gráficamente la comparación de resultados. La variabilidad también es menor en el caso de la actitud muy desfavorable (los datos se encuentran menos dispersos).

Figura 10.8 Ejemplo de interpretación gráfica de las estadísticas descriptivas.



Otro ejemplo de interpretación de los resultados de una medición respecto a una variable es el que ahora se presenta.

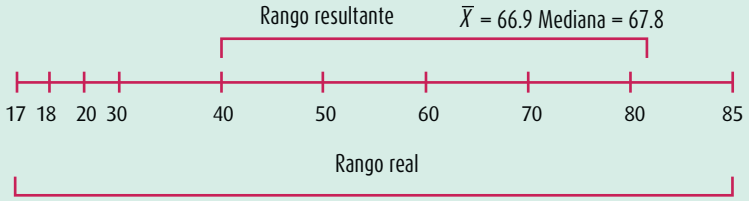
Ejemplo

Hernández-Sampieri y Cortés (1982) aplicaron una prueba de motivación intrínseca sobre la ejecución de una tarea a 60 participantes de un experimento. La escala contenía 17 ítems (con cinco opciones cada uno, uno a cinco) y los resultados fueron los siguientes:[9]

n: 60	Rango: 41	Mínimo: 40	Máximo: 81
Media: 66.9	Mediana: 67.8	Moda: 61	DE: 9.1
Varianza: 83	Curtosis: 0.6	Asimetría: −0.8	EE: 1.18
Sumatoria: 4 013			

¿Qué podríamos decir sobre la motivación intrínseca de los participantes?

El nivel de motivación intrínseca exhibido por los participantes tiende a ser elevado, como lo indican los resultados. El rango real de la escala iba de 17 a 85. El rango resultante para esta investigación varió de 40 a 81. Por tanto, es evidente que los individuos se inclinaron hacia valores elevados en la medida de motivación intrínseca. Además, la media de los participantes es de 66.9 y la mediana de 67.8, lo cual confirma la tendencia de la muestra hacia valores altos de la escala. A pesar de que la dispersión de las puntuaciones de los sujetos es considerable (la desviación estándar es igual a 9.1 y el rango es de 41), esta dispersión se manifiesta en el área más elevada de la escala. Veámoslo gráficamente.



Escala de motivación intrínseca (datos ordinales, supuestos como datos en nivel de intervalo).

En resumen, la tarea resultó intrínsecamente motivante para la mayoría de los participantes; sólo que para algunos resultó muy motivante, para otros, relativamente motivante, y para los demás, medianamente motivante. Esto es, que la tendencia general es hacia valores superiores.

Ahora bien, ¿qué significa un alto nivel de motivación intrínseca exhibido con respecto a una tarea? Implica que la tarea fue percibida como atractiva, interesante, divertida y categorizada como una experiencia agradable. Asimismo, involucra que los individuos, al ejecutarla, derivaron de ella sentimientos de satisfacción, goce y realización personal. Por lo general, quien se encuentra intrínsecamente motivado hacia una labor, disfrutará la ejecución de ésta, ya que obtendrá de la labor *per se* recompensas internas, como sentimientos de logro y autorrealización. Además de ser absorbido por el desarrollo de la tarea y, al tener un buen desempeño, la opinión de sí mismo mejorará o se verá reforzada.

¿Hay alguna otra estadística descriptiva?

Sí, *la asimetría y la curtosis.* Los *polígonos de frecuencia* son *curvas*, por ello se representan como tales (figura 10.9), para que puedan analizarse en términos de probabilidad y visualizar su grado de dispersión. Estos dos elementos resultan esenciales para analizar estas curvas o polígonos de frecuencias.

Asimetría y **curtosis** Estadísticas que se usan para conocer cuánto se parece una distribución a la distribución teórica llamada *curva normal* o campana de Gauss y dónde se concentran las puntuaciones.

La **asimetría** es una estadística necesaria para conocer cuánto se parece nuestra distribución a una distribución teórica llamada *curva normal* (la cual se representa también en la figura 10.9) y constituye un indicador del lado de la curva donde se agrupan las frecuencias. Si es cero (asimetría = 0), la curva o distribución es simétrica. Cuando es positiva, quiere decir que hay más valores agrupados hacia la izquierda de la curva (por debajo de la media). Cuando es negativa, significa que los valores tienden a agruparse hacia la derecha de la curva (por encima de la media) (Hume, 2011; Taylor, 2007a; Salkind, 2006; y Burkhart, 2003).

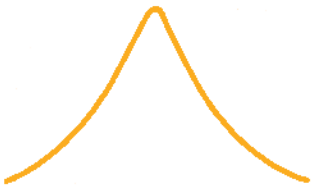
La **curtosis** es un indicador de lo plana o "picuda" que es una curva. Cuando es cero (curtosis = 0), significa que puede tratarse de una *curva normal*. Si es positiva, quiere decir que la curva, la distri-

[9] EE significa "error estándar".

bución o el polígono es más "picudo" o elevado. Si la curtosis es negativa, indica que es más plana la curva (Hume, 2011, Taylor, 2007b, Field, 2006 y Cameron, 2003).

La asimetría y la curtosis requieren al menos un nivel de medición por intervalos. En la figura 10.9 se muestran ejemplos de curvas con su interpretación.

Figura 10.9 Ejemplos de curvas o distribuciones y su interpretación.



Distribución simétrica (asimetría = 0), con curtosis positiva, y una desviación estándar y varianza medias.

Distribución con asimetría negativa, curtosis positiva, y desviación estándar y varianza mayores.

Distribución con asimetría positiva, curtosis negativa, y desviación estándar y varianza considerables.

Distribución con asimetría negativa, curtosis positiva, y desviación estándar y varianza menores.

Distribución simétrica, curtosis positiva, y una desviación estándar y varianza bajas.

Curva normal, curtosis = 0, asimetría = 0, y desviación estándar y varianza promedios.

¿Cómo se traducen las estadísticas descriptivas al inglés?

Algunos programas y paquetes estadísticos computacionales pueden realizar el cálculo de las estadísticas descriptivas, cuyos resultados aparecen junto al nombre respectivo de éstas, muchas veces en inglés.

A continuación se indican las diferentes estadísticas y su equivalente en inglés.

Estadística	Equivalente en inglés
• Moda	• *Mode*
• Mediana	• *Median*
• Media	• *Mean*
• Desviación estándar	• *Standard deviation*
• Varianza	• *Variance*
• Máximo	• *Maximum*
• Mínimo	• *Minimum*
• Rango	• *Range*
• Asimetría	• *Skewness*
• Curtosis	• *Kurtosis*

Nota final

Debe recordarse que en una investigación se obtiene una distribución de frecuencias y se calculan las estadísticas descriptivas para cada *variable*, las que se necesiten de acuerdo con los propósitos de la investigación y los niveles de medición.

Ejemplo

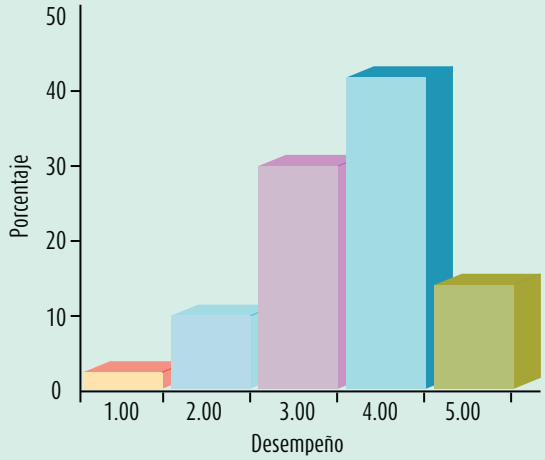
Hernández-Sampieri (2005), en su investigación sobre el clima organizacional, obtuvo las siguientes estadísticas fundamentales de sus variables en una de las muestras:

Variable	*n*	Mínimo	Máximo	Media	Desviación estándar
Moral	390	1.00	5.00	3.3818	0.91905
Dirección	393	1.00	5.00	2.7904	1.08775
Innovación	396	1.00	5.00	3.4621	0.91185
Identificación	383	1.00	5.00	3.6584	0.91283
Comunicación	397	1.00	5.00	3.2519	0.87446
Desempeño	403	1.00	5.00	3.6402	0.86793
Motivación intrínseca	401	2.00	5.00	3.9111	0.73900
Autonomía	395	1.00	5.00	3.2025	0.85466
Satisfacción	399	1.00	5.00	3.7249	0.90591
Liderazgo	392	1.00	5.00	3.4532	1.10019
Visión	391	1.00	5.00	3.7341	0.89206
Recompensas	381	1.00	5.00	2.4528	1.14364

Notas: Todas las variables son compuestas (integradas de varios ítems). La columna *n* representa el número de casos válidos para cada variable. El *n* total de la muestra es de 420, pero como podemos ver en la tabla, el número de casos es distinto en las diferentes variables, porque SPSS elimina de toda la variable los casos que no hayan respondido a un ítem o más reactivos. La variable con mayor promedio es la *motivación intrínseca* y la más baja es *recompensas*.

Posteriormente, obtuvo las tablas y distribuciones de frecuencias de todas sus 12 variables, de las cuales solamente incluimos la variable "desempeño" por cuestiones de espacio.

Desempeño				
	Valores	Frecuencia	Porcentaje válido	Porcentaje acumulado
	1	2	0.5	0.5
	2	35	8.7	9.2
	3	133	33.0	42.2
	4	169	41.9	84.1
	5	64	15.9	100.0
Total *n* = 420 Perdidos = 17		403	100.0	



Para el cálculo de estadísticas descriptivas (tendencia central y dispersión) en SPSS, se sugiere descargar la versión de prueba del sitio de SPSS y consultar el manual respectivo.

SPSS®

Puntuaciones *z*

Las puntuaciones *z* son transformaciones que se pueden hacer a los valores o las puntuaciones obtenidas, con el propósito de analizar su distancia respecto a la media, en unidades de desviación estándar. Una puntuación *z* nos indica la dirección y el grado en que un valor individual obtenido se aleja de la media, en una escala de unidades de desviación estándar. El lector puede conocer más sobre las puntuaciones *z* en el capítulo 8 adicional que puede descargarse del centro de recursos en línea, en: Material complementario → Capítulos → Capítulo 8, "Análisis estadístico: segunda parte".

Centro de recursos en línea

Razones y tasas

Una razón es la relación entre dos categorías. Por ejemplo:

Categorías	Frecuencia
Masculino	60
Femenino	30

La razón de hombres a mujeres es de $\frac{60}{30} = 2$. Es decir, por cada dos hombres hay una mujer.

Una **tasa** es la relación entre el número de casos, frecuencias o eventos de una categoría y el número total de observaciones, multiplicada por un múltiplo de 10, generalmente 100 o 1 000. La fórmula es:

Tasa Relación entre el número de casos de una categoría y el número total de observaciones.

$$\text{Tasa} = \frac{\text{Número de eventos}}{\text{Número total de eventos posibles}} \times 100 \text{ o } 1\,000$$

$$\text{Ejemplo} \quad \frac{\text{Número de nacidos vivos en la ciudad}}{\text{Número de habitantes en la ciudad}} \quad 1\,000$$

$$\text{Tasa de nacidos vivos en Santa Lucía: } \frac{10\,000}{300\,000} \times 1\,000 = 33.33$$

Es decir, hay 33.33 nacidos vivos por cada 1 000 habitantes en Santa Lucía.

Corolario

Hemos analizado descriptivamente los datos por *variable del estudio* y los visualizamos gráficamente. En caso de que alguna distribución resulte ilógica, debemos cuestionarnos si la variable debe ser excluida, sea por errores del instrumento de medición o en la recolección de los datos, ya que la codificación puede ser verificada. Supongamos que en una investigación en empresas, al medir la satisfacción laboral, resulta que 90% se encuentra "sumamente satisfecho" (¿es lógico?); u otro caso sería que, en ingresos anuales el promedio fuera de 15 000 dólares por familia (¿resulta creíble en tal municipio?). La tarea es revisar la información descriptiva de todas las variables y verificar su veracidad.

Asimismo, si nos encontramos un porcentaje alto de valores perdidos (por ejemplo, de 20%),[10] debemos preguntarnos: ¿por qué tantos participantes no respondieron o contestaron erróneamente?, ¿por qué no se tienen registros completos de todos los casos, eventos o unidades de análisis? (como datos de laboratorio en un estudio clínico, mediciones de ciertas reacciones químicas, etcétera).

Ahora, debemos demostrar la confiabilidad y validez de nuestro instrumento, sobre la base de los datos recolectados.

Paso 4: evaluar la confiabilidad o fiabilidad y validez lograda por el instrumento de medición

La confiabilidad se calcula y evalúa para todo el instrumento de medición utilizado, o bien, si se administraron varios instrumentos, se determina para cada uno de ellos. Asimismo, es común que el instrumento contenga varias escalas para diferentes variables o dimensiones, entonces la fiabilidad se establece para cada escala y para el total de escalas (si se pueden sumar, si son aditivas).[11]

Tal como se mencionó en el capítulo 9, existen diversos procedimientos para calcular la confiabilidad de un instrumento conformado por una o varias escalas que miden las variables de la investigación; cuyos ítems, variables de la matriz o indicadores pueden sumarse, promediarse o correlacionarse. Todos utilizan fórmulas que producen coeficientes de fiabilidad que pueden oscilar entre cero y uno, donde recordemos que un coeficiente de cero significa nula confiabilidad y uno representa un máximo de fiabilidad. Cuanto más se acerque el coeficiente a cero (0), mayor error habrá en la medición (Garson, 2013; Franzen, Robbins y Sawicki, 2010; así como Lauriola, 2003). Los coeficientes expresan la intercorrelación (consistencia) entre los distintos ítems, indicadores o componentes de la prueba (Knapp, 2013; Cervantes, 2005; Cortina, 1993; y Carmines y Zeller, 1991).

Los procedimientos más utilizados para determinar la confiabilidad mediante un coeficiente son:[12]

1. *Medida de estabilidad* (confiabilidad por test-retest). En este procedimiento un mismo instrumento de medición se aplica dos o más veces a un mismo grupo de personas o casos, después de cierto periodo. Si la correlación entre los resultados de las diferentes aplicaciones es muy positiva, el instrumento se considera confiable (Rodríguez, 2006a y Krauss y Chen, 2003). Se trata de una especie de diseño de panel. Desde luego, el periodo entre las mediciones es un factor que hay que considerar. Si el periodo es largo y la variable o el contexto son susceptibles de cambios, ello suele confundir la interpretación del coeficiente de fiabilidad obtenido por este procedimiento. Y si

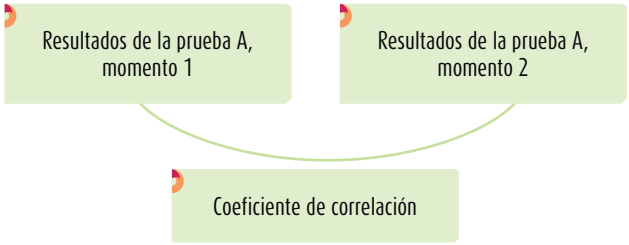
[10] Un porcentaje de valores perdidos (*missing data*) no debe ser mayor de 15%, no es razonable (Creswell, 2005). Cuando tenemos valores perdidos, podemos ignorarlos o sustituirlos por el valor promedio obtenido del total de puntuaciones válidas (esto lo hacen muchos programas de análisis) si así lo deseamos, y puede ser una solución (McKnight *et al.*, 2007). Para profundizar en el tema, véase Leeuw y Hox (2008), Allison (2001) y Enders (2010), en este orden. En las áreas de ingeniería a Latini y Passerini (2003) y en el caso de ciencias de la salud a Allison (2007) y a O'Kelly y Ratitch (2014).

[11] En aparatos o sistemas se verifica la calibración o consistencia entre diversas mediciones.

[12] Un coeficiente de fiabilidad es una medida de la proporción de varianza verdadera en relación con la varianza total observada a través de las puntuaciones o valores resultantes de la administración de un instrumento o protocolo de medición (prueba, escala, calificación de experto, etc.) a una muestra de individuos (Lauriola, 2003). Se aplican más bien a mediciones que involucran respuestas de personas, aunque pueden adaptarse a otros casos que impliquen correlacionar resultados (por ejemplo, a datos de organismos biológicos, protocolos de procesos o aparatos, etcétera).

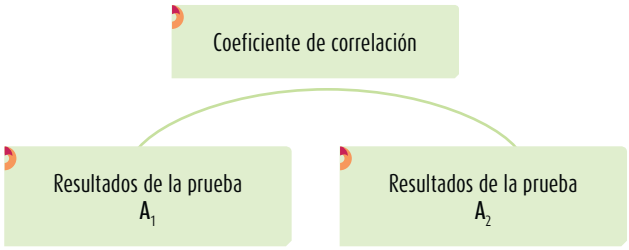
el periodo es corto las personas pueden recordar cómo respondieron en la primera aplicación del instrumento, para aparecer como más consistentes de lo que en realidad son (Bohrnstedt, 1976). El proceso de cálculo con dos aplicaciones se representa en la figura 10.10.

Figura 10.10 Medida de estabilidad.



2. *Método de formas alternativas o paralelas.* En este esquema no se administra el mismo instrumento de medición, sino dos o más versiones equivalentes de éste. Las versiones (casi siempre dos) son similares en contenido, instrucciones, duración y otras características, y se administran a un mismo grupo de personas simultáneamente o dentro de un periodo corto. El instrumento es confiable si la correlación entre los resultados de ambas administraciones es positiva de manera significativa (Rodríguez, 2006b). Los patrones de respuesta deben variar poco entre las aplicaciones. Una variación de este método es el de las formas alternas prueba-posprueba (Creswell, 2005), cuya diferencia reside en que el tiempo que transcurre entre la administración de las versiones es mucho más largo, que es el caso de algunos experimentos. El método se representa en la figura 10.11.

Figura 10.11 Método de formas alternativas o paralelas.



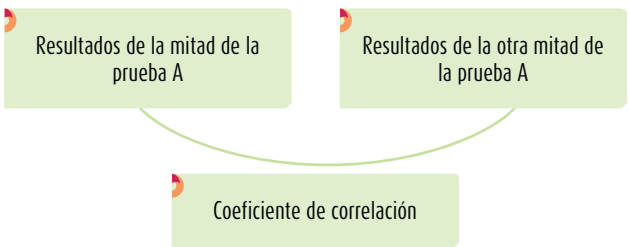
Estos dos métodos (estabilidad y formas alternas) también pueden aplicarse cuando utilizamos dos instrumentos distintos para medir las mismas variables en las unidades de análisis (por ejemplo, dos sistemas para medir propiedades eléctricas o dos protocolos para medir la presión arterial).

3. *Método de mitades partidas (split-halves).* Los procedimientos anteriores requieren cuando menos dos administraciones de la medición en la muestra. En cambio, el **método de mitades partidas** necesita sólo una aplicación de la medición. Específicamente, el conjunto total de ítems o reactivos se divide en dos mitades equivalentes y se comparan las puntuaciones o resultados de ambas. Si el instrumento es confiable, las puntuaciones de las dos mitades deben estar muy correlacionadas (Rodríguez, 2006c y McKelvie, 2003). Un individuo con baja puntuación en una mitad tenderá a mostrar también una baja puntuación en la otra mitad. El procedimiento se diagrama en la figura 10.12.

Métodos de mitades partidas y de consistencia interna Se aplican a instrumentos que implican medidas compuestas o escalas, es decir, están constituidas por varios ítems, indicadores o mediciones.

4. *Medidas de coherencia o consistencia interna*. Éstos son coeficientes que estiman la confiabilidad: *a*) *el alfa de Cronbach* (desarrollado por J.L. Cronbach) y *b*) *los coeficientes KR-20 y KR-21* de Kuder y Richardson (1937). El método de cálculo de éstos requiere una sola administración del instrumento de medición. Su ventaja reside en que no es necesario dividir en dos mitades a los ítems del instrumento, simplemente se aplica la medición y se calcula el coeficiente. La mayoría de los programas estadísticos como SPSS y Minitab los determinan y solamente deben interpretarse.

Figura 10.12 Método de mitades partidas.



Respecto a la interpretación de los distintos coeficientes mencionados cabe señalar que no hay una regla que indique "a partir de este valor no hay fiabilidad del instrumento". Más bien, el investigador calcula su valor, lo declara y lo somete a escrutinio de los usuarios del estudio u otros investigadores, explicitando el método utilizado (Chen y Krauss, 2003; McKelvie, 2003; Lauriola, 2003; y Carmines y Zeller, 1991). Algunos autores consideran que el coeficiente debe estar entre 0.70 y 0.90 (Tavakol y Dennick, 2011; DeVellis, 2003; Streiner, 2003; Nunnally y Bernstein, 1994; Petterson, 1994). Nunnally (1987) por encima de 0.80. Lauriola (2003) sugiere un valor mínimo de 0.70 para la comparación entre grupos y 0.90 para escalas. Garson (2013) establece que 0.60 es aceptable para

propósitos exploratorios y 0.70 para fines confirmatorios, resultando 0.80 "bueno" en un alcance explicativo. Ahora bien, también un coeficiente mayor de 0.90 puede implicar redundancia de ítems o indicadores y la necesidad de reducir el instrumento (Tavakol y Dennick, 2011).

Con respecto a los métodos basados en coeficientes de correlación, usted se formará una idea más clara después de revisar el apartado de correlación que se presenta más adelante en este capítulo. Pero hay una consideración importante que hacer ahora. El coeficiente que elijamos para determinar la confiabilidad debe ser apropiado al nivel de medición de la escala de nuestra variable (por ejemplo, si la escala de mi variable es por intervalos, puedo utilizar el coeficiente de correlación de Pearson; pero si es ordinal podré usar el coeficiente de Spearman o de Kendall; y si es nominal, otros coeficientes). El *alfa* de *Cronbach* trabaja con variables de intervalos o de razón, KR-20 y KR-21 con ítems dicotómicos (por ejemplo: sí-no)[13] y $\rho xx'$ con reactivos tricotómicos (Knapp, 2013; Alkharusi, 2010; Vittengl, White, McGovern y Morton, 2006; y Feldt, 2005). Además, existen otros coeficientes como el alfa estratificado, la confiabilidad máxima, los coeficientes de Raju, Kristof, Angoff-Feldt, Feldt-Gilmer, Guttman λ_2, λ_4 maximizado y el análisis de Hoyt.

El cálculo del coeficiente *alfa* y algunas consideraciones sobre los factores que lo afectan se incluyen en el capítulo 8 adicional: "Análisis estadístico: segunda parte", que se encuentra en "Material complementario".

Centro de recursos en línea

Con la finalidad de comprender mejor los métodos para determinar la confiabilidad, véase la tabla 10.10.

Tabla 10.10 Aspectos básicos de los métodos para determinar la confiabilidad de instrumentos aplicados a personas

Método	Número de veces en que el instrumento es administrado	Número de versiones diferentes del instrumento	Número de participantes que proveen los datos	Inquietud o pregunta que contesta
Estabilidad (*test-retest*)	Dos veces en tiempos distintos	Una versión	Cada participante responde al instrumento dos veces.	¿Responden los individuos de una manera similar a un instrumento si se les administra dos veces?
Formas alternas	Dos veces al mismo tiempo o con una diferencia de tiempo muy corta	Dos versiones diferentes, pero equivalentes	Cada participante responde a cada versión del instrumento.	Cuando dos versiones de un instrumento son similares, ¿hay convergencia o divergencia en las respuestas a ambas versiones?
Formas alternas y prueba-posprueba	Dos veces en tiempos distintos	Dos versiones diferentes, pero equivalentes	Cada participante responde a cada versión del instrumento.	Cuando dos versiones de un instrumento son similares, ¿hay convergencia o divergencia en las respuestas a ambas versiones?
Mitades partidas	Una vez	Una fragmentada en dos partes equivalentes	Cada participante responde a la única versión.	¿Son las puntuaciones de una mitad del instrumento similares a las obtenidas en la otra mitad?
Medidas de consistencia interna (alfa y KR-20 y 21)	Una vez	Una versión	Cada participante responde a la única versión.	¿Las respuestas a los ítems del instrumento son coherentes?

[13] Estos dos coeficientes se usan en el método de "mitades partidas", aunque —como señalan Babbie (2012) y Creswell (2005)— se confía en la mitad de la información del instrumento, por lo que conviene agregar el cálculo de "profecía" Spearman-Brown.

Asimismo, en la tabla 10.11 se presentan ejemplos de estudios con su respectiva confiabilidad.

Tabla 10.11 Ejemplos de confiabilidad

Investigación	Instrumento	Métodos de cálculo y resultados	Comentario
Evaluación de los conocimientos, opiniones, experiencias y acciones en torno al abuso sexual infantil (Kolko *et al.*, 1987).	Escala cognitiva de nueve ítems para infantes en edades preescolares y primeros grados básicos.	Coherencia interna alfa = 0.34.	Confiabilidad extremadamente baja que demuestra incongruencia, atribuida por los autores a lo corto de la escala (pocos ítems).
Desarrollo y validación de una escala autoaplicable para medir la satisfacción sexual en adultos varones y mujeres de México (Álvarez Gayou, Honold y Millán, 2005).	Un inventario para medir la satisfacción sexual que está integrado por 29 reactivos y fue administrado a una muestra de 760 personas, de ambos géneros, cuyas edades fluctuaron entre los 16 y 65 años.	La confiabilidad del inventario establecida por medio de una prueba alfa de Cronbach fue de 0.92.	El valor α indica una fiabilidad sumamente elevada.
Validación de un instrumento para medir la cultura empresarial en función del clima organizacional y vincular empíricamente ambos constructos (Hernández-Sampieri, Méndez y Contreras, 2013).	Cuestionario estandarizado que mide el clima organizacional en función del Modelo de los Valores en Competencia de Quinn y Rohrbaugh, a través de escalas tipo Likert con cuatro opciones de respuesta: dos positivas y dos negativas.	El coeficiente alfa-Cronbach obtenido resultó igual a 0.95 (con 95 ítems). La muestra estuvo conformada por 1 424 empleados de 12 empresas (972 casos válidos completos).	Confiabilidad muy elevada. No hay redundancia de ítems pues se midieron 17 variables del clima organizacional (asociadas entre sí). Los coeficientes para las escalas variaron entre 0.60 y 0.90 (y uno sólo de 0.53).
Actitudes hacia el matrimonio: integración y sus resultados en las relaciones personales (Riggio y Weiser, 2008).	Escalas del Modelo de Inversión (IMS), las cuales a partir de 37 reactivos (cada uno con 9 categorías) miden la entrega, la inversión psicológica y la satisfacción con respecto a una relación romántica actual.	Los coeficientes alfa resultantes de aplicar las escalas a 400 universitarios fueron: 0.94 para entrega y satisfacción, y 0.88 para inversión psicológica.	Coeficientes muy considerables para entrega y satisfacción, y bastante aceptable para inversión.
Fukuda, Saklofske, Tamaoka y Lim, 2012.	Prueba WLEIS de 16 reactivos que mide la inteligencia emocional en adultos jóvenes en cuatro dimensiones: valoración de las propias emociones, valoración de las emociones de los demás, uso de las emociones y control o regulación de las emociones.	Los coeficientes resultantes de aplicar el WLEIS (n=161) fueron por subescala o dimensión: valoración de las propias emociones: 0.80, valoración de las emociones de los demás: 0.74, uso de las emociones: 0.74 y regulación de las emociones: 0.83.	Coeficientes aceptables dentro de los parámetros normales, particularmente porque las escalas tienen pocos ítems.

Otro caso es el ya comentado de Núñez (2001) y su instrumento para medir el sentido de vida, cuya fiabilidad fue de 0.96 en su tercera versión con 99 ítems.

Como podemos observar en la tabla 10.11, cuanta más información se proporcione sobre la confiabilidad, el lector se forma una idea más clara sobre su cálculo y las condiciones en que se demostró. Es indispensable incluir las dimensiones de la variable medida, el tamaño de muestra y el método utilizado. Una cuestión importante es que los coeficientes son sensibles al número de ítems o reactivos: entre más agreguemos, el valor del coeficiente tenderá a ser más elevado.

Además de estimar un coeficiente de correlación o un coeficiente de coherencia entre los ítems del instrumento, es conveniente calcular la correlación ítem-escala completa. Ésta representa la vinculación de cada reactivo con toda la escala. Habrá tantas correlaciones como ítems contenga el instrumento. Corbetta (2003, p. 237) lo ejemplifica adecuadamente de la siguiente manera: si estamos midiendo el autoritarismo, es lógico pensar que quien alcanza altas puntuaciones en esta variable en toda la escala (es muy autoritaria), habrá de tener puntuaciones elevadas en todos los ítems que la conforman. Pero si uno de los reactivos sistemáticamente (en un número considerable de individuos) presenta valores contradictorios con respecto a la escala total, podemos concluir que ese ítem no fun-

ciona correctamente (contradice a los demás reactivos). Los ítems que alcancen coeficientes de correlación bajos con la escala tal vez deban analizarse y, eventualmente, eliminarse.

Asimismo, cada uno de los reactivos puede ser evaluado en su capacidad de discriminación mediante la prueba *t* de Student (paramétrica). Se consideran dos grupos, el primero integrado por 25% de los casos con los puntajes más altos obtenidos en el ítem y el otro grupo compuesto por 25% de los casos con los puntajes más bajos. Los ítems cuya prueba no resulte significativa serán reconsiderados.

Los conceptos estadísticos aquí vertidos (por ejemplo, correlación y prueba *t*) tendrán mayor sentido, una vez que se revisen más ampliamente, lo cual se hará más adelante en este capítulo.

Para determinar la confiabilidad usando los programas estadísticos no olvide consultar los respectivos manuales, descargándolos del centro de recursos en línea.

Validez

Vimos en el capítulo anterior que la evidencia sobre la validez del contenido se obtiene mediante las opiniones de expertos y al asegurarse de que las dimensiones medidas por el instrumento sean representativas del universo o dominio de dimensiones de las variables de interés (a veces mediante un muestreo aleatorio simple). La evidencia de la validez de criterio se produce al correlacionar las puntuaciones de los participantes, obtenidas por medio del instrumento, con sus valores logrados en el criterio. Recordemos que una correlación implica asociar puntuaciones obtenidas por la muestra en dos o más variables.

Por ejemplo, Núñez (2001), además de aplicar su instrumento sobre el sentido de vida, administró otras dos pruebas que teóricamente miden variables similares: el PIL (Propósito de Vida) y el Logo-test de Elizabeth Lukas. El coeficiente de correlación de Pearson entre el instrumento diseñado y el PIL fue de 0.541, valor que se considera moderado. El coeficiente de correlación *rho* de Spearman fue igual a 0.42 entre el Logo-test y su prueba, lo cual indica que los tres instrumentos no miden la misma variable, pero sí conceptos relacionados.

La evidencia de la validez de constructo se obtiene mediante el análisis de factores. Tal método nos indica cuántas dimensiones integran a una variable y qué ítems conforman cada dimensión. Los reactivos que no pertenezcan a una dimensión, quiere decir que están "aislados" y no miden lo mismo que los demás ítems, por tanto, deben eliminarse. Es un método que tradicionalmente se ha considerado complejo, por los cálculos estadísticos implicados, pero que es relativamente sencillo de interpretar y como los cálculos hoy en día los realiza la computadora, está al alcance de cualquier persona que se inicie dentro de la investigación. Este método se revisa —con ejemplos reales— en el capítulo 8 adicional del centro de recursos en línea: "Análisis estadístico: segunda parte".

Para cada escala, una vez que se determina la confiabilidad (de 0 a 1) y se muestra la evidencia sobre la validez, si algunos ítems son problemáticos (no discriminan, no se vinculan a otros ítems, van en sentido contrario a toda la escala, no miden lo mismo, etc.), se eliminan de los cálculos (pero en el reporte de la investigación, se indica cuáles fueron descartados, las razones de ello y cómo alteran los resultados); posteriormente se vuelve a realizar el análisis descriptivo (distribución de frecuencias, medidas de tendencia central y de variabilidad, etcétera).

En el centro de recursos → Material complementario → Ejemplos → Ejemplo 4, "Diseño de una escala autoaplicable para la evaluación de la satisfacción sexual en hombres y mujeres mexicanos" (Álvarez Gayou, Honold y Millán, 2005), se presenta la validación de un instrumento que muestra todos los elementos para ello, paso por paso. Incluye la generación de redes semánticas. Su abordaje es desde el punto de vista de la salud y con propiedad científica. Se recomienda descargarlo y revisarlo.

¿Hasta aquí llegamos?

Cuando el estudio tiene una finalidad puramente exploratoria o descriptiva, debemos interrogarnos: ¿podemos establecer relaciones entre variables? En caso de una respuesta positiva, es factible seguir con la estadística inferencial; pero si dudamos o el alcance se limitó a explorar y describir, el trabajo de análisis concluye y debemos comenzar a preparar el reporte de la investigación.

Paso 5: analizar mediante pruebas estadísticas las hipótesis planteadas (análisis estadístico inferencial)

En este paso se analizan las hipótesis a la luz de pruebas estadísticas que a continuación detallamos.

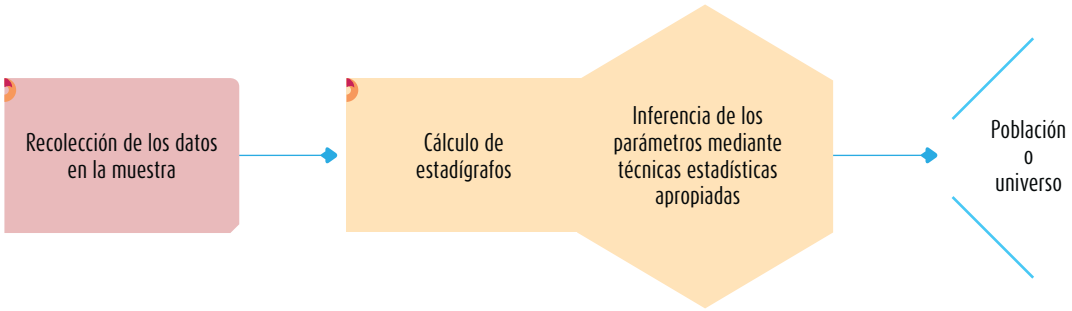
Estadística inferencial: de la muestra a la población

3 y 4

Con frecuencia, el propósito de la investigación va más allá de describir las distribuciones de las variables: se pretende probar hipótesis y generalizar los resultados obtenidos en la muestra a la población o universo. Los datos casi siempre se recolectan de una muestra y sus resultados estadísticos se denominan *estadígrafos*; la media o la desviación estándar de la distribución de una muestra son estadígrafos. A las estadísticas de la población se les conoce como *parámetros*. Éstos no son calculados, porque no se recolectan datos de toda la población, pero pueden ser inferidos de los estadígrafos, de ahí el nombre de **estadística inferencial**. El procedimiento de esta naturaleza de la estadística se esquematiza en la figura 10.13.

Estadística inferencial Estadística para probar hipótesis y estimar parámetros.

Figura 10.13 Procedimiento de la estadística inferencial.



Entonces, la estadística inferencial se utiliza fundamentalmente para dos procedimientos vinculados (O'Leary, 2014; Punch, 2014; Babbie, 2012; Wiersma y Jurs, 2008; Waterman, 2007; Kulikowich y Edwards, 2006; y Maxim, 2003):

a) *Probar hipótesis poblacionales*
b) *Estimar parámetros*

En este capítulo comentaremos la prueba de hipótesis, que se efectúa dependiendo del tipo de hipótesis de que se trate. Existen pruebas estadísticas para diferentes clases de hipótesis como iremos viendo.

La inferencia de los parámetros depende de que hayamos elegido una muestra probabilística con un tamaño que asegure un nivel de significancia o significación adecuado (Jarman, 2013; Lindsay, 2009; y Moriceau, 2009). En el centro de recursos encontrará un ejemplo de inferencia sobre la hipótesis de la media poblacional, en: Material Complementario → Capítulos → Capítulo 8, "Análisis estadístico: segunda parte".

Centro de recursos en línea

¿En qué consiste la prueba de hipótesis?

3

Una hipótesis en el contexto de la estadística inferencial es una proposición respecto de uno o varios parámetros, y lo que el investigador hace por medio de la prueba de hipótesis es determinar si la hipótesis poblacional es congruente con los datos obtenidos en la muestra (Wilcox, 2012; Gordon, 2010; Wiersma y Jurs, 2008; y Stockburger, 2006).

Una hipótesis se retiene como un valor aceptable del parámetro, si es consistente con los datos. Si no lo es, se rechaza (pero los datos no se descartan). Para comprender lo que es la prueba de hipó-

tesis en la estadística inferencial es necesario revisar los conceptos de distribución muestral[14] y nivel de significancia.[15]

¿Qué es una distribución muestral?

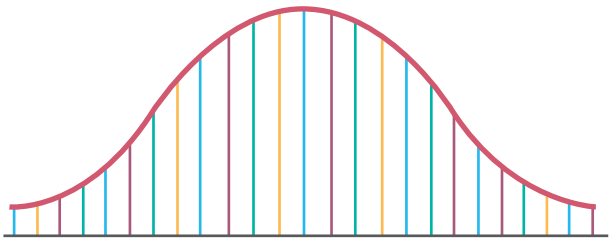
Distribución muestral Conjunto de valores sobre una estadística calculada de todas las muestras posibles de una población.

Una **distribución muestral** es un conjunto de valores sobre una estadística calculada de todas las muestras posibles de determinado tamaño de una población (Bond, 2007a). Las distribuciones muestrales de medias son probablemente las más conocidas. Expliquemos este concepto con un ejemplo. Supongamos que nuestro universo son los automovilistas de una ciudad y deseamos averiguar cuánto tiempo pasan diariamente manejando ("al volante"). De este universo podría extraerse una muestra representativa. Vamos a suponer que el tamaño adecuado de muestra es de 512 automovilistas ($n = 512$). Del mismo universo se podrían extraer diferentes muestras, cada una con 512 personas.

3

Teóricamente, incluso podría elegirse al azar una, dos, tres, cuatro muestras, y las veces que fuera necesario hacerlo, hasta agotar todas las muestras posibles de 512 automovilistas de esa ciudad (todos los individuos serían seleccionados en varias muestras). En cada muestra se obtendría una media del tiempo que pasan los automovilistas manejando. Tendríamos pues, una gran cantidad de medias, tantas como las muestras extraídas ($\overline{X}_1$, $\overline{X}_2$, $\overline{X}_3$, $\overline{X}_4$, $\overline{X}_5$, ... $\overline{X}_k$). Y con éstas elaboraríamos una distribución de medias. Habría muestras que, en promedio, pasaran más tiempo "al volante" que otras. Este concepto se representa en la figura 10.14.

Figura 10.14 Distribución muestral de medias.



Son medias ($\overline{X}$) no se trata de puntuaciones. Cada media representaría una muestra.

Si calculáramos la media de todas las medias de las muestras, prácticamente obtendríamos el valor de la media poblacional. De hecho, casi nunca se obtiene la distribución muestral (la distribución de las medias de todas las muestras posibles). Es más bien un concepto teórico definido por la estadística para los investigadores. Lo que comúnmente hacemos es extraer una sola muestra.

En el ejemplo de los automovilistas, sólo una de las líneas verticales de la distribución muestral presentada en la figura 10.14 es la media obtenida para nuestra única muestra seleccionada de 512 personas. Y la pregunta es: ¿nuestra media calculada se encuentra cerca de la media de la distribución muestral?, debido a que si está cerca podremos tener una estimación precisa de la media poblacional (el parámetro poblacional es prácticamente el mismo que el de la distribución muestral). Esto se expresa en el *teorema central del límite*:

> Si una población (no necesariamente normal) tiene de media m y de desviación estándar s, la distribución de las medias en el muestreo aleatorio realizado en esta población tiende, al aumentar n, a una distribución normal de media m y desviación estándar $\frac{s}{\sqrt{n}}$, donde n es el tamaño de muestra.

El teorema especifica que la distribución muestral tiene una media igual a la de la población, una varianza igual a la varianza de la población dividida entre el tamaño de muestra (su desviación estándar es $\frac{\sigma}{\sqrt{n}}$ y se distribuye normalmente). La desviación estándar (s) es un parámetro normalmente desconocido, aunque es posible estimarlo por la desviación estándar de la muestra. Asimismo, en el capítulo 8 se dijo que cuando las muestras están constituidas por 100 o más elementos tienden a presentar **distribuciones normales** y esto sirve para el propósito de hacer estadística inferencial. La "normalidad" de la distribución en muestras grandes no obedece a la normalidad de la distribución de una población. La distribución de diversas variables a veces es "normal" y en ocasiones está lejos de serlo. Sin embargo, la normalidad no debe confundirse con probabilidad. Mientras lo primero es necesario para efectuar ciertas pruebas estadísticas, lo segundo es requisito indispensable para

Distribución normal Distribución en forma de campana que se logra con muestras de 100 o más unidades muestrales y que es útil y necesaria cuando se hacen inferencias estadísticas.

[14] Distribución muestral y distribución de una muestra son conceptos diferentes: la última es resultado de los datos de nuestra investigación y es por variable.

[15] El término *significancia* es un anglicismo, por lo que diversos autores sugieren mejor utilizar "significación" o "significatividad" (por ejemplo: Korniejczuk, 2012).

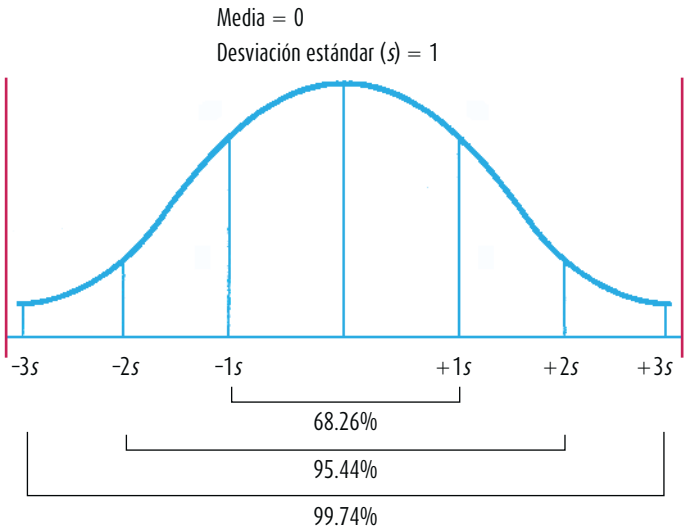
www.elosopanda.com | jamespoetrodriguez.com

hacer inferencias correctas sobre una población. El concepto de *distribución normal* es importante otra vez y se ofrece una breve explicación en la figura 10.15.

Una gran cantidad de los fenómenos del comportamiento humano se manifiestan de la siguiente forma: la mayoría de las puntuaciones se concentran en el centro de la distribución, en tanto que en los extremos encontramos sólo algunas puntuaciones (Fu, 2007). Por ejemplo, la inteligencia: hay pocas personas muy inteligentes (genios), pero también hay pocas personas con muy baja inteligencia. La mayoría de los seres humanos somos medianamente inteligentes. Esto podría representarse así:

Figura 10.15 Concepto de curva o distribución normal.

Debido a ello, se creó un modelo de probabilidad llamado curva normal o distribución normal. Como todo modelo es una distribución conceptual que difícilmente se presenta en la realidad tal cual, pero sí se presentan aproximaciones a éste. La curva normal tiene la siguiente configuración:



68.26% del área de la curva normal es cubierta entre −1*s* y +1*s*, 95.44% del área de esta curva es cubierta entre −2*s* y +2*s* y 99.74% se cubre con −3*s* y +3*s*.

Las principales características de la distribución normal son:

1. Es *unimodal*, una sola moda.
2. La *asimetría es cero*. La mitad de la curva es exactamente igual a la otra mitad. La distancia entre la media y −3*s* es la misma que la distancia entre la media y +3*s*.
3. *Es una función* particular entre desviaciones con respecto a la media de una distribución y la probabilidad de que éstas ocurran.
4. La base está dada en unidades de desviación estándar (puntuaciones *z*), destacando las puntuaciones −1*s*, −2*s*, −3*s*, +1*s*, +2*s* y +3*s* (que equivalen respectivamente a −1.00*z*, −2.00*z*, −3.00*z*, +1.00*z*, +2.00*z*, +3.00*z*). Las distancias entre puntuaciones *z* representan áreas bajo la curva. De hecho, la distribución de puntuaciones *z* es la curva normal.
5. Es *mesocúrtica* (curtosis de cero).
6. La *media*, la *mediana* y la *moda* coinciden en el mismo punto (el centro).

Nivel de significancia Nivel de la probabilidad de equivocarse y que fija de manera *a priori* el investigador.

¿Qué es el nivel de significancia o significación?

Wiersma y Jurs (2008) ofrecen una explicación sencilla del concepto, en la cual nos basaremos para analizar su significado. La probabilidad de que un evento ocurra oscila entre cero (0) y uno (1), donde cero implica la imposibilidad de ocurrencia y uno la certeza de que el fenómeno ocurra. Al lanzar al aire una moneda no cargada, la probabilidad de que salga "cruz" es de 0.50 y la probabilidad de que la moneda caiga en "cara" también es de 0.50. Con un dado, la probabilidad de obtener cualquiera de sus caras al lanzarlo es de 1/6 = 0.1667. La suma de posibilidades siempre es de uno.

Aplicando el concepto de probabilidad a la distribución muestral, tomaremos el área de ésta como 1.00; en consecuencia, cualquier área comprendida entre dos puntos de la distribución corresponderá a la probabilidad de la distribución. Para probar hipótesis inferenciales respecto a la media, el investigador debe evaluar si es alta o baja la probabilidad de que la media de la muestra esté cerca de la media de la distribución muestral. Si es baja, el investigador dudará de generalizar a la población. Si es alta, el investigador podrá hacer generalizaciones. Es aquí donde entra el nivel de significancia o nivel alfa (α),[16] el cual es un nivel de la probabilidad de equivocarse y se fija antes de probar hipótesis inferenciales.

Este concepto fue esbozado en el capítulo 8 con un ejemplo coloquial, pero lo volvemos a recordar: si fuera a apostar en las carreras de caballos y tuviera 95% de probabilidades de atinarle al ganador, contra sólo 5% de perder, ¿apostaría? Obviamente sí, siempre y cuando le aseguraran ese 95% en favor.

Pues bien, algo parecido hace el investigador. Obtiene una estadística en una muestra (por ejemplo, la media) y analiza qué porcentaje tiene de confianza en que dicha estadística se acerque al valor de la distribución muestral (que es el valor de la población o el parámetro). Busca un alto porcentaje de certeza, una probabilidad elevada para estar tranquilo, porque sabe que tal vez haya error de muestreo y, aunque la evidencia parece mostrar una aparente "cercanía" entre el valor calculado en la muestra y el parámetro, tal "cercanía" puede no ser real o deberse a errores en la selección de la muestra.

¿Con qué porcentaje de confianza el investigador generaliza, para suponer que tal cercanía es real y no por un error de muestreo? Existen dos niveles convenidos en las ciencias:

a) *El nivel de significancia de 0.05*, el cual implica que el investigador tiene 95% de seguridad para generalizar sin equivocarse y sólo 5% en contra. En términos de probabilidad, 0.95 y 0.05, respectivamente; ambos suman la unidad. Este nivel es el más común en ciencias sociales.
b) *El nivel de significancia de 0.01*, el cual implica que el investigador tiene 99% en su favor y 1% en contra (0.99 y 0.01 = 1.00) para generalizar sin temor. Muy utilizado cuando las generalizaciones implican riesgos vitales para las personas (pruebas de vacunas, medicamentos, arneses de aviones, resistencia de materiales de construcción al fuego o el peso, etcétera).

A veces el nivel de significancia o significación puede ser todavía más riguroso, por ejemplo, 0.001, 0.00001, 0.00000001 (Liao, 2003), pero al menos debe ser de 0.05. No se acepta un nivel de 0.06 (94% a favor de la generalización confiable), porque se busca hacer ciencia lo más exacta posible.[17]

Tal nivel es un valor de certeza que el investigador fija *a priori*, respecto a no equivocarse (Capraro, 2006). Cuando uno lee en un reporte de investigación que los resultados fueron significativos al nivel de 0.05 ($p < 0.05$), indica lo que se comentó: que existe 5% de posibilidad de error al aceptar la hipótesis, correlación o valor obtenido al aplicar una prueba estadística; o 5% de riesgo de que se rechace una hipótesis nula cuando era verdadera (Babbie, 2012 y Mertens, 2010). Volveremos más adelante sobre este punto.

[16] No confundir con el coeficiente alfa de Cronbach, para determinar la confiabilidad.
[17] El nivel de significancia mínimo aceptable es definido por las asociaciones científicas correspondientes al ramo o área en la cual se investiga, incluyendo comités editoriales de revistas académicas.

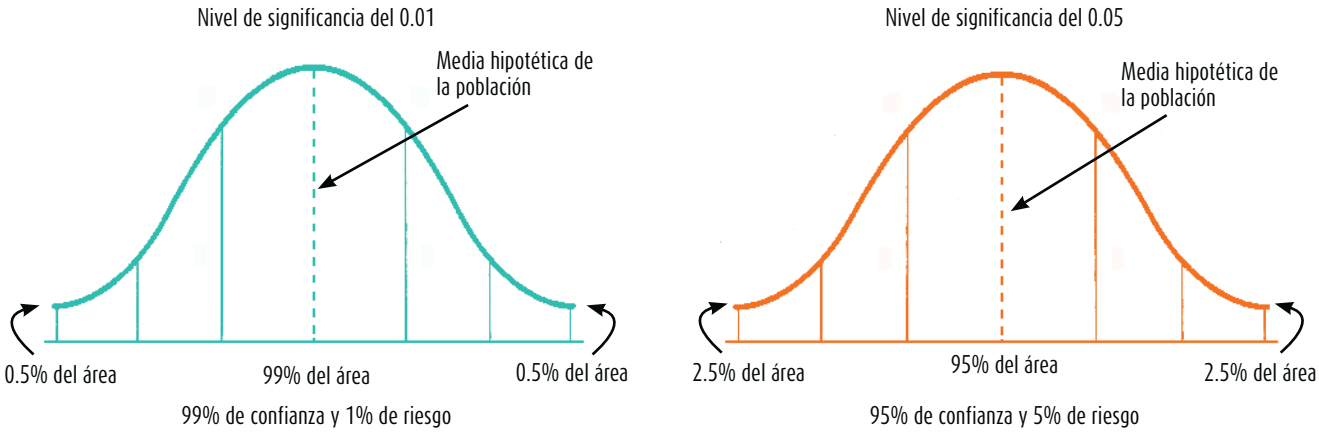
www.elosopanda.com | jamespoetrodriguez.com

¿Cómo se relacionan la distribución muestral y el nivel de significancia?

El *nivel de significancia o significación* se expresa en términos de probabilidad (0.05 y 0.01) y la *distribución muestral* también como probabilidad (el área total de ésta como 1.00). Pues bien, para ver si existe o no confianza al generalizar acudimos a la distribución muestral, con una probabilidad adecuada para la investigación. Dicho nivel lo tomamos como un área bajo la distribución muestral, como se observa en la figura 10.16, y depende de si elegimos un nivel de 0.05 o de 0.01. Es decir, que nuestro valor estimado en la muestra no se encuentre en el área de riesgo y estemos lejos del valor de la distribución muestral, que insistimos es muy cercano al de la población.

Así, el nivel de significación representa áreas de riesgo o confianza en la distribución muestral.

Figura 10.16 Niveles de significancia o significación en la distribución muestral.



Notas:

1. Podemos expresarlo en proporciones (0.025, 0.95 y 0.025, respectivamente) o porcentajes como está en la gráfica.
2. Tanto 99% como 95% representan las áreas de confianza de que nuestra estimación se localiza dentro de ellas. La primera al nivel del 0.01 y la segunda al nivel de 0.05. El área de riesgo en el primer caso es de 1% (0.5 + 0.5 = 1%) y en el segundo de 5% (2.5% + 2.5% = 5%) sumando ambos extremos, porque en nuestra estimación de la media poblacional podríamos pasarnos (error) hacia valores más altos o bajos.

¿Se pueden cometer errores al probar hipótesis y realizar estadística inferencial?

3 y 4

Nunca estaremos completamente seguros de nuestra estimación. Trabajamos con altos niveles de confianza o seguridad, pero, aunque el riesgo es mínimo, podría cometerse un error. *Los resultados posibles al probar hipótesis son*:[18]

1. Aceptar una hipótesis verdadera (decisión *correcta*).
2. Rechazar una hipótesis falsa (decisión *correcta*).
3. Aceptar una hipótesis falsa (conocido como *error* del *Tipo II o error beta*).
4. Rechazar una hipótesis verdadera (conocido como *error* del *Tipo I o error alfa*).

Ambos tipos de error son indeseables; sin embargo, puede *reducirse sustancialmente la posibilidad* de que se presenten mediante:

a) *Muestras probabilísticas representativas.*
b) *Inspección cuidadosa de los datos.*
c) *Selección de las pruebas estadísticas apropiadas.*
d) *Mayor conocimiento de la población.*

[18] Yaremko *et al.* (2013), Cozby y Bates (2012), Ravid (2011), Mertens (2010), Buskirk (2008) y Wiersma y Jurs (2008).

Prueba de hipótesis

4

Hay dos tipos de análisis estadísticos que pueden realizarse para probar hipótesis: los *análisis paramétricos* y los *no paramétricos.* Cada tipo posee sus características y presuposiciones que lo sustentan; la elección de qué clase de análisis efectuar depende de los supuestos. De igual forma, cabe destacar que en una misma investigación es posible llevar a cabo análisis paramétricos para algunas hipótesis y variables, y análisis no paramétricos para otras. Asimismo, como vimos, los análisis a realizar dependen del planteamiento, tipo de hipótesis y el nivel de medición de las variables que las conforman.

Análisis paramétricos

3 y 4

Para realizar análisis paramétricos debe partirse de los siguientes supuestos:[19]

1. La *distribución poblacional de la variable dependiente es normal*: el universo tiene una distribución normal.
2. El *nivel de medición* de las variables es *por intervalos o razón*.
3. Cuando *dos o más poblaciones son estudiadas, tienen una varianza homogénea*: las poblaciones en cuestión poseen una dispersión similar en sus distribuciones.

Ciertamente estos criterios son tal vez demasiado rigurosos y algunos investigadores sólo basan sus análisis en el tipo de hipótesis y los niveles de medición de las variables. Esto queda a juicio del lector. En la investigación académica y cuando quien la realiza es una persona experimentada, sí debe solicitársele tal rigor.

¿Cuáles son los métodos o las pruebas estadísticas paramétricas más utilizados?

Existen diversas pruebas paramétricas, pero las más utilizadas son:

- Coeficiente de correlación de Pearson y regresión lineal.
- Prueba *t*.
- Prueba de contraste de la diferencia de proporciones.
- Análisis de varianza unidireccional (ANOVA en un sentido).
- Análisis de varianza factorial (ANOVA).
- Análisis de covarianza (ANCOVA).

Centro de recursos en línea

Algunos de estos métodos se tratan aquí en este capítulo y otros se explican en el capítulo 8 adicional, "Análisis estadístico: segunda parte", que puede descargarse del centro de recursos en línea de la obra.

Cada prueba obedece a un tipo de hipótesis de investigación e hipótesis estadística distinta. Las hipótesis estadísticas se comentan en el capítulo 8 del centro de recursos en línea.

¿Qué es el coeficiente de correlación de Pearson?

Es una prueba estadística para analizar la relación entre dos variables medidas en un nivel por intervalos o de razón. Se le conoce también como "coeficiente producto-momento".

Se simboliza: *r*

Hipótesis a probar: correlacional, del tipo de "a mayor *X*, mayor *Y*", "a mayor *X*, menor *Y*", "altos valores en *X* están asociados con altos valores en *Y*", "altos valores en *X* se asocian con bajos valores de *Y*". La hipótesis de investigación señala que la correlación es significativa.

Variables: dos. La prueba en sí no considera a una como independiente y a otra como dependiente, ya que no evalúa la causalidad. La noción de causa-efecto (independiente-dependiente) es posible establecerla teóricamente, pero la prueba no asume dicha causalidad.

[19] O'Leary (2014), Ryan (2013), Babbie (2012), Martin y Bridgmon (2012), Kantor y Kershaw (2010), y Wiersma y Jurs (2008).

El coeficiente de correlación de Pearson se calcula a partir de las puntuaciones obtenidas en una muestra en dos variables. Se relacionan las puntuaciones recolectadas de una variable con las puntuaciones obtenidas de la otra, con los mismos participantes o casos (The SAGE Glossary of the Social and Behavioral Sciences, 2009g; Bagiella, 2007; Onwuegbuzie, Daniel y Leech, 2006a).

Nivel de medición de las variables: intervalos o razón.

Interpretación: el coeficiente *r* de Pearson *puede variar de −1.00 a +1.00,* donde:

−1.00 = *correlación negativa perfecta*. ("A mayor *X*, menor *Y*", de manera proporcional. Es decir, cada vez que *X* aumenta una unidad, *Y* disminuye siempre una cantidad constante). Esto también se aplica "a menor *X*, mayor *Y*".

−0.90 = Correlación negativa muy fuerte.
−0.75 = Correlación negativa considerable.
−0.50 = Correlación negativa media.
−0.25 = Correlación negativa débil.
−0.10 = Correlación negativa muy débil.
0.00 = No existe correlación alguna entre las variables.
+0.10 = Correlación positiva muy débil.
+0.25 = Correlación positiva débil.
+0.50 = Correlación positiva media.
+0.75 = Correlación positiva considerable.
+0.90 = Correlación positiva muy fuerte.
+1.00 = *Correlación positiva perfecta* ("A mayor *X*, mayor *Y*" o "a menor *X*, menor *Y*", de manera proporcional. Cada vez que *X* aumenta, *Y* aumenta siempre una cantidad constante).

El *signo indica la dirección de la correlación* (positiva o negativa); y *el valor numérico, la magnitud de la correlación.* Los principales programas computacionales de análisis estadístico indican si el coeficiente es o no significativo de la siguiente manera:

$r = 0.7831$	(valor del coeficiente)
s o $P = 0.001$	(significancia)
$N = 625$	(número de casos correlacionados)

Si *s* o *P* es menor del valor 0.05, se dice que el coeficiente es *significativo* en el nivel de 0.05 (95% de confianza en que la correlación sea verdadera y 5% de probabilidad de error). Si es menor a 0.01, el coeficiente es *significativo* al nivel de 0.01 (99% de confianza de que la correlación sea verdadera y 1% de probabilidad de error).

O bien, otros programas como IBM SPSS® presentan los coeficientes de correlación en una tabla, donde las filas o columnas son las variables asociadas y se señala con asterisco(s) el nivel de significancia: un asterisco (*) implica que el coeficiente es significativo al nivel del 0.05 y dos asteriscos (**) que es significativo al nivel del 0.01. Esto podemos verlo en el ejemplo de la tabla 10.12:

Tabla 10.12 Correlaciones entre moral y dirección

Correlaciones			
		Moral	**Dirección**
Moral	Correlación de Pearson	1	0.557**
	Sig. (bilateral)		0.000
	N	362	335
Dirección	Correlación de Pearson	0.557**	1
	Sig. (bilateral)	0.000	
	N	335	373

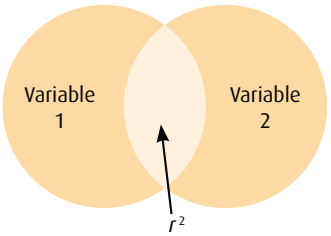
** La correlación es significativa al nivel 0.01 (bilateral, en ambos sentidos entre las variables).

Obsérvese que se correlacionan dos variables: "moral" y "dirección", aunque la correlación aparece dos veces, porque es una tabla que hace todas las comparaciones posibles entre las variables y al hacerlo, genera un eje diagonal (representado por las correlaciones de las variables contra ellas mismas —"moral" con "moral" y "dirección" con "dirección", que carece de sentido porque son las mismas puntuaciones, por eso es perfecta—), y por encima de ese eje aparecen todos los coeficientes, y se repiten por debajo del eje. La correlación es de 0.557 y es significativa en el nivel del 0.000 (menor del 0.01). *N* representa el número de casos correlacionados.

Una correlación de Pearson puede ser significativa, pero si es menor a 0.30 resulta débil, aunque de cualquier manera ayuda a explicar el vínculo entre las variables. Si queremos asociar la *presión arterial* y *el peso* de un grupo de pacientes, la *solubilidad del gas* con la *temperatura* (en ingeniería petrolera) y la *inversión en publicidad* y las *ventas*, es útil este coeficiente.

Consideraciones: cuando el coeficiente *r* de Pearson se eleva al cuadrado (r^2), se obtiene el coeficiente de determinación y el resultado indica la *varianza de factores comunes.* Esto es, el porcentaje de la variación de una variable debido a la variación de la otra variable y viceversa (o cuánto explica o determina una variable la variación de la otra). Veámoslo gráficamente en la figura 10.17.

Figura 10.17 Varianza de factores comunes.



Por ejemplo, si la correlación entre "productividad" y "asistencia al trabajo" es de 0.80.

$$r = 0.80$$
$$r^2 = 0.64$$

"La productividad" constituye a, o explica, 64% de la variación de "la asistencia al trabajo".

"La asistencia al trabajo" explica 64% de "la productividad".

Si *r* es 0.72 y consecuentemente $r^2 = 0.52$, quiere decir que poco más de la mitad de la variabilidad de un constructo o variable está explicada por la otra.

Ejemplo

Hi:	"A mayor motivación intrínseca, mayor productividad".
Resultado:	$r = 0.721$ s o $P = 0.0001$
Interpretación:	se acepta la hipótesis de investigación en el nivel de 0.01. La correlación entre la motivación intrínseca y la productividad es considerable y positiva.
Hi:	"a mayor ingreso, mayor motivación intrínseca".
Resultado:	$r = 0.214$ s o $P = 0.081$
Interpretación:	se acepta la hipótesis nula. El coeficiente no es significativo: 0.081 es mayor que 0.05; recordemos que 0.05 es el nivel mínimo para aceptar la hipótesis.

Nota precautoria: recuerde lo referente a correlaciones espurias que se comentaron en el capítulo 5, "Definición del alcance de la investigación por realizar".

Creswell (2005) señala que un coeficiente de determinación (r^2) entre 0.66 y 0.85 ofrece una buena predicción de una variable respecto de la otra variable; y por encima de 0.85 implica que ambas variables miden casi el mismo concepto subyacente, son "cercanamente" un constructo semejante.

El coeficiente de correlación de Pearson es útil para relaciones lineales, como lo veremos en la regresión lineal, pero no para relaciones curvilineales; en este caso o cuando las variables son ordinales, se suele usar la *rho* de Spearman (ρ) (Onwuegbuzie, Daniel y Leech, 2006b).

Cuando queremos correlacionar simultáneamente más de dos variables, por ejemplo: motivación, satisfacción en el trabajo, moral y autonomía; o como lo hicieron Wood *et al.* (2009) con productividad del médico (medida en unidades de valor relativo McGladrey, MRVU), tiempo médico/paciente, satisfacción y confianza del paciente, se utiliza el coeficiente de correlación múltiple o *R*, el cual se revisa en el capítulo adicional 8 "Análisis estadístico: segunda parte", del centro de recursos.

Centro de recursos en línea

Para el cálculo del coeficiente de correlación de Pearson mediante IBM SPSS®, no olvide consultar el manual respectivo. En Minitab en Estadísticas → Estadísticas básicas.

¿Qué es la regresión lineal?

3

Es un modelo estadístico para estimar el efecto de una variable sobre otra. Está asociado con el coeficiente *r* de Pearson. Brinda la oportunidad de predecir las puntuaciones de una variable a partir de las puntuaciones de la otra variable. Entre mayor sea la correlación entre las variables (covariación), mayor capacidad de predicción.

Hipótesis: correlacionales y causales.

Variables: dos. Una se considera como independiente y otra como dependiente. Pero, para poder hacerlo, debe tenerse un sólido sustento teórico.

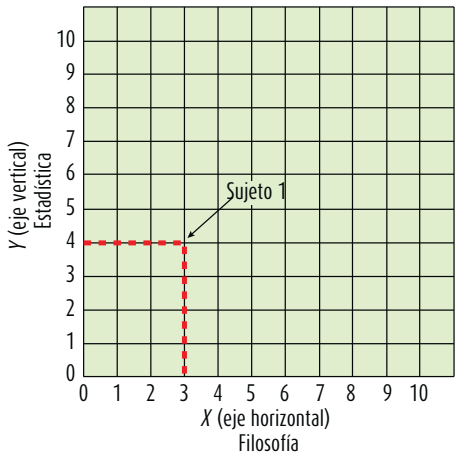
Nivel de medición de las variables: intervalos o razón.

Procedimiento e interpretación: la regresión lineal se determina con base en el diagrama de dispersión. Éste consiste en una gráfica donde se relacionan las puntuaciones de una muestra en dos variables (Martin y Bridgmon, 2012; Bednarczyk y McNutt, 2007; Harrington, 2007; Daniel, Onwuegbuzie y Leech, 2006; y Wood y Park, 2003). Veámoslo con un ejemplo sencillo de ocho casos. Una variable es la calificación en Filosofía y la otra variable es la calificación en Estadística; ambas medidas, hipotéticamente, de 0 a 10.

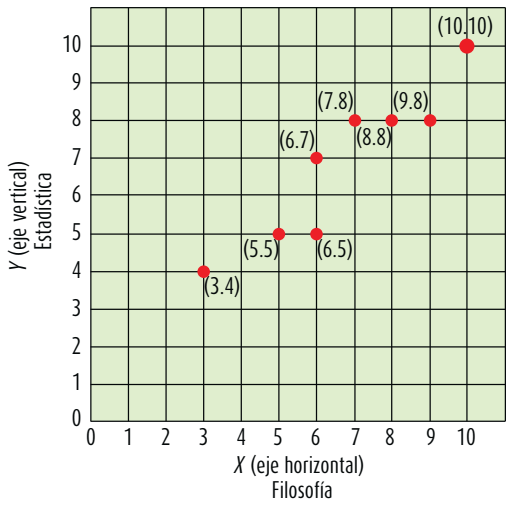
	Puntuaciones	
Sujetos	Filosofía (*X*)	Estadística (*Y*)
1	3	4
2	8	8
3	9	8
4	6	5
5	10	10
6	7	8
7	6	7
8	5	5

Figura 10.18 Ejemplo de gráficas de dispersión.

El *diagrama de dispersión* se construye graficando cada par de puntuaciones en un espacio o plano bidimensional. Sujeto "1" tuvo 3 en *X* (filosofía) y 4 en *Y* (estadística):

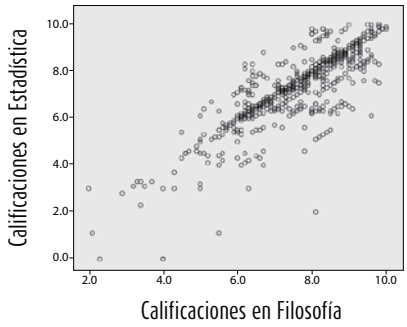


Así se grafican todos los pares:



(*continúa*)

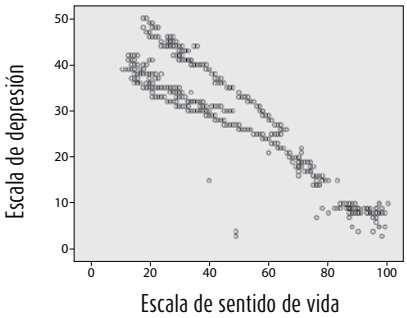
Figura 10.18 (*continuación*)



Los *diagramas de dispersión* son una manera de visualizar gráficamente una correlación. Por ejemplo:

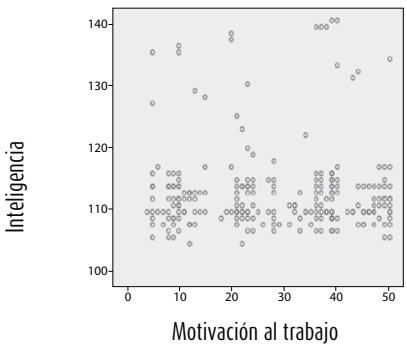
Si aplicáramos los exámenes de Filosofía y Estadística (escala de 0 a 10 en ambas mediciones) a 775 alumnos y obtuviéramos el siguiente resultado: $r = 0.814$** (significativa al nivel del .01). La correlación es considerablemente positiva y el diagrama de dispersión sería el siguiente:[20]

La tendencia es ascendente, altas puntuaciones en *Y*, altas puntuaciones en *X* (mejores calificaciones en Estadística están asociadas con mejores calificaciones en Filosofía).

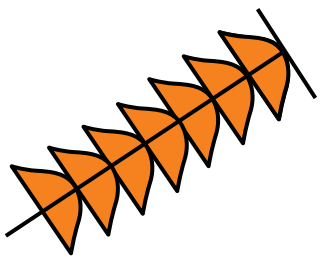


En cambio, si administráramos una prueba sobre la "depresión" (escala de 0 a 50) y una que mida el "sentido de vida" (0 a 100) y el resultado fuera: -0.926** (significativa al nivel del .01). La correlación es sumamente negativa y el diagrama de dispersión sería el siguiente:

La tendencia es descendente, altas puntuaciones en *depresión* se encuentran vinculadas con bajas en *sentido de vida*, y viceversa.



En el caso de que dos variables no estén correlacionadas, por ejemplo: $r = .006$ (no significativa) (digamos entre "inteligencia" −90 a 140− y "motivación al trabajo" −0 a 50−). El diagrama de dispersión no tiene ninguna tendencia:



Así, cada punto representa un caso y un resultado de la intersección de las puntuaciones en ambas variables. El diagrama de dispersión puede ser resumido a una línea, si hay tendencia.

Conociendo la línea y la tendencia, podemos predecir los valores de una variable conociendo los de la otra variable (Shapiro, 2008).

Esta línea es la recta de regresión y se expresa mediante la *ecuación de regresión lineal*:

$$Y = a + bX$$

en donde *Y* es un *valor de la variable dependiente* que se desea predecir, *a* es la *ordenada* en el origen (intersección) y *b* la *pendiente* o inclinación, *X* es el valor que fijamos en la variable independiente o *predictora*.

Los programas de análisis estadístico que incluyen la *regresión lineal*, proporcionan los datos de *a* y *b*.

a o intercepción (*intercept*) y *b* o pendiente (*slope)*

[20] Estos diagramas fueron visualizados a través de SPSS, versión 15.

Para predecir un valor de *Y*, se sustituyen los valores correspondientes en la ecuación.

Ejemplo

a (intercepción) = 1.2

b (pendiente) = 0.8

Entonces podemos hacer la predicción: ¿a un valor de 7 en Filosofía qué valor le corresponde en Estadística?

$$Y = \underset{a}{1.2} + \underset{b}{(0.8)}\ \underset{X}{(7)}$$

$$Y = 6.8$$

Predecimos que a un valor de 7 en *X* le corresponderá un valor de 6.8 en *Y*.

Ejemplo

Regresión lineal

Hi: "La autonomía laboral es una variable que predice la motivación intrínseca en el trabajo. Ambas variables están relacionadas".
Las dos variables fueron medidas en una escala de 1 a 5.

Resultado: a (intercepción) = 0.42
b (pendiente) = 0.65

Interpretación: cuando *X* (autonomía) es 1, la predicción estimada de *Y* es 1.07; cuando *X* es 2, la predicción estimada de *Y* es 1.72; cuando *X* es 3, *Y* será 2.37; cuando *X* es 4, *Y* será 3.02; y cuando *X* es 5, *Y* será 3.67.

$$Y = a + bX$$
$$1.07 = 0.42 + 0.65\ (1)$$
$$1.72 = 0.42 + 0.65\ (2)$$
$$2.37 = 0.42 + 0.65\ (3)$$
$$3.02 = 0.42 + 0.65\ (4)$$
$$3.67 = 0.42 + 0.65\ (5)$$

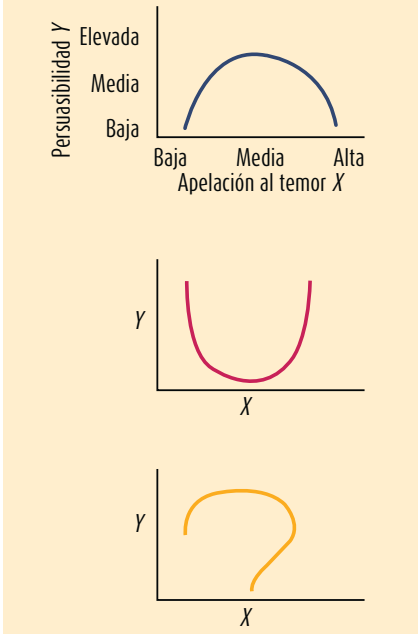
Consideraciones: la regresión lineal es útil con *relaciones lineales,* no con *relaciones curvilineales* (Graham, 2013; Bates y Watts, 2007; y Little, 2003). Porque como señalan León y Montero (2003, p. 191) es un error atribuir a la relación causal una covariación exclusivamente lineal: a mayores valores en la variable independiente, mayores valores en la dependiente. Existen muchas relaciones de causa-efecto que no son lineales, como por ejemplo: la vinculación entre ansiedad y rendimiento. Cierto grado de ansiedad ayuda a conseguir mejores resultados en un examen o la práctica de un deporte; pero, por encima de determinado nivel (nerviosismo extremo), la ejecución empeora. También, la dosis de un medicamento puede tener una relación no lineal con el efecto esperado (cierta dosis no lograrlo por insuficiente y demasiada dosis, provocar otros efectos indeseados con consecuencias muy negativas). Asimismo, determinadas reacciones químicas necesitan una temperatura específica (ni más, ni menos) y lo mismo con la cantidad de riego para una parcela donde se experimenta con cultivos específicos. En la figura 10.19 se muestran ejemplos de estas relaciones.

Las *relaciones curvilineales* son aquellas en las cuales la tendencia varía, por ejemplo: primero es ascendente y luego descendente, o viceversa.

Se ha demostrado que una estrategia persuasiva con niveles altos de apelación al temor, por ejemplo, un comercial televisivo muy dramático, provoca una escasa persuasión, lo mismo que una estrategia persuasiva con niveles muy bajos de apelación al temor.

La estrategia persuasiva más adecuada es la que utiliza niveles medios de apelación al temor. Esta relación es curvilineal; véase figura 10.19.

Figura 10.19 Ejemplos de relaciones curvilineales.



En la práctica, los estudiantes no deben preocuparse por graficar los diagramas de dispersión. Esto lo hace el programa respectivo (SPSS®, Minitab u otro).

¿Qué es la prueba *t*?

Es una prueba estadística para evaluar si dos grupos difieren entre sí de manera significativa respecto a sus medias en una variable.

Se simboliza: *t*.

Hipótesis: de diferencia entre dos grupos. La hipótesis de investigación propone que los grupos difieren entre sí de manera significativa y la hipótesis nula plantea que los grupos no difieren significativamente.

Los grupos pueden ser dos plantas comparadas en su productividad, dos escuelas contrastadas en los resultados a un examen, dos clases de materiales de construcción cotejados en su rendimiento, dos medicamentos comparados en su efecto, etcétera.

Variables: la comparación se realiza sobre una variable (regularmente y de manera teórica: dependiente). Si hay diferentes variables, se efectuarán varias pruebas *t* (una por cada variable), y la razón que motiva la creación de los grupos puede ser una variable independiente. Por ejemplo, un experimento con dos grupos, donde a uno se le aplica el estímulo experimental y al otro no, es de control.

Nivel de medición de la variable de comparación: intervalos o razón.

Cálculo e interpretación: el valor *t* es calculado por el programa estadístico.[21] Los programas, por ejemplo SPSS, arrojan una tabla con varios resultados, de los cuales los más necesarios para interpretar son el valor *t* y su significancia. Veamos un ejemplo y la interpretación.[22]

Ejemplo

Hi: "Los varones le atribuyen mayor importancia al atractivo físico en sus relaciones heterosexuales que las mujeres".

Ho: "Los varones *no* le atribuyen mayor importancia al atractivo físico en sus relaciones heterosexuales que las mujeres".

La variable *atractivo físico* fue medida a través de una escala que varía de 0 a 18. El grupo de mujeres estuvo constituido por 119 personas y el de hombres por 128 (variable que origina el contraste: género). Los resultados fueron:

$\bar{X}_1$ (mujeres) = 12

$\bar{X}_2$ (hombres) = 15

Valor t = 6.698 (significancia menor de 0.01)

n_1 = 119 mujeres

n_2 = 128 hombres

Grados de libertad = 245

Conclusión: se acepta la hipótesis de investigación y se rechaza la hipótesis nula.
Si el valor *t* hubiera sido de 1.05 y no significativo, se aceptaría la hipótesis nula.

La prueba *t* se basa en una distribución muestral o poblacional de diferencia de medias conocida como la distribución *t* de Student que se identifica por los grados de libertad, los cuales constituyen el número de maneras en que los datos pueden variar libremente. Son determinantes, ya que nos indican qué valor debemos esperar de *t*, dependiendo del tamaño de los grupos que se comparan. *Cuanto mayor número de grados de libertad se tengan, la distribución t de Student se acercará más a ser una distribución normal* y usualmente, si los grados de libertad exceden los 120, la distribución nor-

Centro de recursos en línea

[21] Para quienes se interesen en las fórmulas para calcular manualmente el valor de la prueba *t*, se encuentran en el capítulo 8, "Análisis estadístico: segunda parte", en Material complementario → Capítulos.

[22] Se evita aquí la discusión sobre si las percepciones pueden medirse en un nivel de intervalos u ordinal. El ejemplo intenta atraer la atención de los alumnos. Desde luego, si el profesor las considera ordinales, puede cambiar el ejemplo por otro que considere pertinente en su campo.

mal se utiliza como una aproximación adecuada de la distribución *t* de Student (Babbie, 2012, Wiersma y Jurs, 2008; y Godby, 2007).

Los grados de libertad se calculan con la fórmula siguiente, en la que n_1 y n_2 son el tamaño de los grupos que se comparan:

$$gl = (n_1 + n_2) - 2$$

Vogt (1999) señala que los grados de libertad indican cuántos casos fueron usados para calcular un valor estadístico en particular.

Hernández-Sampieri *et al.* (2010) realizaron un análisis mediante la prueba *t* con poco menos de medio millón de alumnos de una institución pública, con la finalidad de comparar el desempeño entre mujeres y hombres respecto al promedio general de la carrera, el valor obtenido fue de 22.802, significancia = 0.000 (menor al 0.01). El promedio de los estudiantes fue de 6.58 ($n = 302\ 272$) y el de las estudiantes de 7.11 ($n = 193\ 436$). Ante la interrogante: ¿se observaron diferencias en el desempeño académico por género? Se puede decir que las mujeres obtienen mayor promedio que los hombres en una diferencia de 0.53 puntos, la cual es significativa al nivel del 0.01.

Así como el caso anterior, esta prueba se utiliza frecuentemente para hacer contrastes por género. Por ejemplo, los efectos de administrar un medicamento (digamos, en la presión arterial u otra variable) o seguir una dieta (en la disminución del peso corporal). Buunk, Castro, Zurriaga y González (2011) llevaron a cabo un estudio en España y Argentina para comparar si los hombres son más celosos que las mujeres ante la presencia de un *rival* en cuatro dimensiones o características: atractivo físico, dominación física, poder social y atributos sociales comunitarios. Resultaron significativos al nivel del 0.01, todos los valores "t"[23] de las dos primeras (77.98 en España y 121.89 en Argentina para el atractivo físico, siendo la media de celos mayor en las mujeres; y 21.67 en España y 42.38 en Argentina para dominación física, siendo la media de celos mayor en los hombres). Esto significa que los hombres experimentaron más celos que las mujeres cuando su rival es físicamente más dominante. En cambio, las mujeres experimentaron más celos que los hombres cuando su rival era más atractiva (a los hombres les preocupa la "musculatura de sus rivales" y a las mujeres "el atractivo físico"). La *n* fue de 388 españoles y 444 argentinos de ambos géneros.

Consideraciones: La prueba *t* se utiliza para comparar los resultados de una preprueba con los resultados de una posprueba en un contexto experimental. Se comparan las medias y las varianzas del grupo en dos momentos diferentes: $\bar{X}_1 \times \bar{X}_2$. O bien, para comparar las prepruebas o pospruebas de dos grupos que participan en un experimento:

G_1 $\quad \bar{X}_1$

G_2 $\quad \bar{X}_2 \quad \overset{t}{\bigcirc}$ son las pospruebas

Cuando el valor *t* se calcula mediante un paquete estadístico computacional, la significancia se proporciona como parte de los resultados y debe ser menor a 0.05 o 0.01, lo cual depende del nivel de confianza seleccionado (regularmente se ofrece el resultado en dos versiones, según sea el caso, si se asumen o no varianzas iguales).[24] Lo más importante es visualizar el valor *t* y su significancia; véase la tabla 10.13.

Para solicitar en SPSS la prueba *t*, no olvide consultar el manual "Introducción al IBM SPSS®" que se puede descargar del centro de recursos. En Minitab este método se encuentra en: Estadísticas → Estadísticas básicas. En STATS® se denomina: diferencia de dos medias (Difference-Two Means) y simplemente se colocan número de casos o respuestas en cada grupo, medias y desviaciones estándar de los grupos, y automáticamente se calcula el valor *t* y el nivel de significancia expresado en porcentaje.

Centro de recursos en línea

[23] Ellos lo expresan como valores "F", que prácticamente es lo mismo, implicando también análisis de las varianzas.

[24] Cuando se incluyen participantes diferentes en los grupos del experimento, el diseño se considera de "grupos independientes" (León y Montero, 2003) y no se asumen varianzas iguales.

Tabla 10.13 Elementos fundamentales para interpretar los resultados de una prueba *t*

Estadísticos de grupo

F3	Género	N	Media	Desviación tip.		Error tip. de la media
	Masculino	86	3.69	1.043		0.113
	Femenino	88	3.84	1.071		0.114

Prueba de muestras independientes

		Prueba de Levene para la igualdad de varianzas		Prueba *t* para la igualdad de medias						
									95% intervalo de confianza para la diferencia	
		F	Sig.	*t*	gl	Sig. (bilateral)	Diferencia de medias	Error tip. de la diferencia	inferior	superior
F3	Se han asumido varianzas iguales	0.001	0.970	−0.966	172	0.335	−0.15	0.160	−0.471	0.162
	No se han asumido varianzas iguales			−0.966	171.98	0.335	−0.15	0.160	−0.471	0.162

Valor "F" diferencia entre las varianzas de los grupos (dispersión de los datos)

Valor "t"

Significancia: no es menor al 0.05, mucho menos al 0.01: No hay diferencias entre los grupos en la variable de contraste

¿Qué es el tamaño del efecto?

2 Al comparar grupos, en este caso con la prueba *t*, es importante determinar el tamaño del efecto, que es una medida de la "fuerza" de la diferencia de las medias u otros valores considerados (Creswell, 2013a; Alhija y Levy, 2009; y Cortina, 2003). Resulta ser una medida en unidades de desviación estándar.

¿Cómo se calcula? El tamaño del efecto es justo la diferencia estandarizada entre las medias de los dos grupos. En otras palabras:

$$\text{Tamaño total del efecto} = \frac{\text{Media del grupo 1} - \text{Media del grupo 2}}{\text{Desviación estándar sopesada}}$$

La desviación estándar sopesada es la estimación reunida de la desviación estándar de ambos grupos, basada en la premisa que cualquier diferencia entre sus desviaciones es solamente debida a la variación del muestreo (Rodríguez, 2006 y Creswell, 2005).

La desviación estándar sopesada (denominador en la fórmula) se calcula así:

$$\sqrt{\frac{(N_E - 1)SD_E^2 + (N_C - 1)SD_C^2}{N_E + N_C - 2}}$$

Donde N_E y N_C son el tamaño de los grupos (grados de libertad), respectivamente; en tanto que, SD_E y SD_C son sus desviaciones estándares.

Ejemplo

17.9 − 15.2/3.3 = 0.82 (interpretación: las medias varían menos de una desviación estándar, una respecto de la otra).

Ejemplo

28.5 − 37.5/4.1 = −2.19 (los promedios varían más de dos desviaciones estándar uno sobre otro).

¿Qué es la prueba de diferencia de proporciones?

Es una prueba estadística para analizar si dos proporciones o porcentajes difieren significativamente entre sí.

Hipótesis: de diferencia de proporciones en dos grupos.

Variable: la comparación se realiza sobre una variable. Si hay varias, se efectuará una prueba de diferencia de proporciones por variable.

Nivel de medición de la variable de comparación: cualquier nivel, incluso por intervalos o razón, pero siempre expresados en proporciones o porcentajes.

Procedimiento e interpretación: este análisis puede realizarse muy fácilmente en el programa STATS®, subprograma: Diferencia de dos proporciones (Difference-Two Percentages). Se colocan el número de casos y el porcentaje obtenido para cada grupo y se calcula. Eso es todo. No se necesita de fórmulas y tablas como se hacía anteriormente.

Ejemplo

Hi: "El porcentaje de liberales en la ciudad de Arualm es mayor que en Linderbuck".

En STATS® colocamos los datos que se nos requiere:

Inputs

Grupo uno

Número de respondientes en grupo uno

410

Porcentaje medido en grupo uno

55

Grupo dos

Número de respondientes en grupo uno

301

Porcentaje medido en grupo uno

48

Calculamos **Calculate** y se obtienen los resultados:

Results

Probabilidad de diferencia significativa

94.52%

Valor Z

1.92

Como no se alcanza una significancia de 95% (porque STATS®, al contrario de SPSS® o Minitab, proporciona el porcentaje de ésta a favor), aceptamos la hipótesis nula y rechazamos la de investigación.

Con esta prueba podemos analizar, por ejemplo, si el porcentaje de mujeres con cáncer de mama es significativamente diferente en dos comunidades, si el porcentaje de errores en la producción de arneses automotrices es significativamente distinto en dos plantas, si el porcentaje de reprobados es significativamente desigual entre los alumnos de bachillerato del turno matutino y del vespertino, etc. Desde luego, no es necesario que los grupos por comparar tengan el mismo número de unidades, casos, respondientes o equivalentes (n), salvo las consideraciones de muestreo hechas previamente (tamaño mínimo de grupos).

¿Qué es el análisis de varianza unidireccional o de un factor? (ANOVA *one-way*)

3 Es una prueba estadística para analizar si más de dos grupos difieren significativamente entre sí en cuanto a sus medias y varianzas. La *prueba t* se aplica para *dos grupos* y el *análisis de varianza unidireccional* se usa para *tres, cuatro o más grupos.* Aunque con dos grupos se puede utilizar también.

Hipótesis: de diferencia entre más de dos grupos. La hipótesis de investigación propone que los grupos difieren significativamente entre sí y la hipótesis nula propone que los grupos no difieren significativamente.

Variables: una variable independiente y una variable dependiente.

Nivel de medición de las variables: la variable independiente es categórica y la dependiente es por intervalos o razón.

El hecho de que la variable independiente sea categórica significa que es posible formar grupos diferentes (Martin y Bridgmon, 2012 y Lazar, 2006). Puede ser una variable nominal, ordinal, por intervalos o de razón (pero en estos últimos dos casos la variable debe reducirse a categorías).

Por ejemplo:[25]

- Religión (católica, cristiana, protestante, judía, musulmana, budista, etc.) (puede compararse la satisfacción de los grupos con su religión o el grado de espiritualidad: Soto, 2014).
- Nivel socioeconómico (muy alto, alto, medio, bajo y muy bajo) (contrastarse su lealtad a la marca).
- Antigüedad del empleado en la empresa (de cero a un año, más de un año a cinco años, más de cinco años a 10, más de 10 años a 20 y más de 20 años) (cotejarse su productividad).
- Estadios del cáncer de próstata (I, II, III y IV) (comparar su grado de depresión).
- Obesidad y peso: peso insuficiente, normopeso, sobrepeso, obesidad en grados (I, II, III y IV —extrema—) (cotejar sus niveles de glucosa y presión arterial).
- Giro de la empresa: comercial, industrial y de servicios (comparar los efectos de una medida fiscal en su nivel de tributación).
- Tipo de concreto premezclado (estándar, de fraguado rápido, reforzado con fibras, autocompactante, poroso, antibacteriano, etc.) (contrastar su resistencia).

Análisis de varianza Prueba estadística para analizar si más de dos grupos difieren entre sí de manera significativa en sus medias y varianzas.

Interpretación: el **análisis de varianza** unidireccional produce un valor conocido como *F* o *razón F*, que se basa en una distribución muestral, conocida como *distribución F*, la cual es otro miembro de la familia de distribuciones muestrales. La *razón F* compara las variaciones en las puntuaciones debidas a dos diferentes fuentes: variaciones entre los grupos que se comparan y variaciones dentro de los grupos. Si el valor *F* es significativo implica que los grupos difieren entre sí en sus promedios (Zhang, 2013; The SAGE Glossary of the Social and Behavioral Sciences, 2009h;, Klugkist, 2008; Field, 2006a y 2006b; y Norpoth, 2003). Entonces se acepta la hipótesis de investigación y se rechaza la nula.[26] A continuación se presenta un ejemplo de un estudio en el que el análisis apropiado es el de varianza.

Ejemplo

Hi: "Los niños que se expongan a contenidos de elevada violencia televisiva exhibirán una conducta más agresiva en sus juegos, respecto de los niños que se expongan a contenidos de mediana o baja violencia televisada".

Ho: "Los niños que se expongan a contenidos de elevada violencia televisiva no exhibirán una conducta más agresiva en sus juegos, respecto de los niños que se expongan a contenidos de mediana o baja violencia televisada".

[25] Ejemplos sencillos, simplemente para que el lector que comienza con estos temas tenga una idea de las múltiples aplicaciones del ANOVA. Su profesor puede proporcionar diversas aplicaciones a su área.

[26] El sustento y explicación del análisis de varianza unidireccional que en las primeras cuatro ediciones se incluía en esta parte, ahora la puede encontrar el lector en el centro de recursos: Material Complementario → Capítulos → Capítulo 8, "Análisis estadístico: segunda parte". Le recomiendo descargar y revisar este capítulo.

Centro de recursos en línea

La variable independiente es el grado de exposición a la violencia televisada y la variable dependiente es la agresividad exhibida en los juegos, medida por el número de conductas agresivas observadas (nivel de medición por intervalos).

Para probar la hipótesis se diseña un experimento con cuatro grupos:

G_1X_1 (elevada violencia)
G_2X_2 (mediana violencia)
G_3X_3 (baja violencia)
G_4 – (conducta prosocial)

Número de actos agresivos

En cada grupo hay 25 niños.

La razón *F* fue de 9.89 y resultó significativa en el nivel de 0.05: se acepta la hipótesis de investigación. La diferencia entre las medias de los grupos es admitida, el contenido sumamente violento tiene un efecto sobre la conducta agresiva de los niños en sus juegos. El estímulo experimental tuvo un impacto. Esto se corrobora comparando las medias de las pospruebas de los cuatro grupos, porque el análisis de varianza unidireccional sólo nos señala si la diferencia entre las medias y las distribuciones de los grupos es o no significativa; pero no nos indica en favor de qué grupos lo es. Es posible hacer esto último al visualizar los promedios y compararlos con las distribuciones de sus grupos. Y si adicionalmente queremos cotejar cada par de medias ($\bar{X}_1$ con $\bar{X}_2$, $\bar{X}_1$ con $\bar{X}_3$, $\bar{X}_2$ con $\bar{X}_3$, etc.) y determinar con exactitud dónde están las diferencias significativas, podemos aplicar un contraste posterior, con el cálculo de una prueba *t* para cada par de medias; o bien, por medio de algunas estadísticas que suelen ser parte de los análisis efectuados en los paquetes estadísticos computacionales.

Tales estadísticas se incluyen en la tabla 10.14.

Tabla 10.14 Principales estadísticas para comparaciones posteriores (*post hoc*) en el ANOVA unidireccional o de un factor[27]

Nombre	Siglas
• Diferencia menos significativa	DMS
• Prueba *F* de Ryan-Einot-Gabriel-Welsch	R-E-G-W F
• Prueba de rango de Ryan-Einot-Gabriel-Welsch	R-E-G-W Q
• Prueba de Tukey	
• Otras: Waller-Duncan, T2 de Tamhane, T3 de Dunnett, Games-Howell, *C* de Dunett, Bonferroni, Sidak, Gabriel, Hochberg, Scheffé...	

Ejemplo

Supongamos que por medio de una escala de Likert (1-5)[28] medimos la actitud que tienen hacia el entrenador del equipo de fútbol de una ciudad, las tres porras o grupos de aficionados permanentes: la Ultra, la Central y la de Veteranos. Y queremos analizar si difieren significativamente entre sí. Realizamos el análisis de varianza y los resultados son los que se muestran en la tabla 10.15 con los elementos que suelen incluir los programas de análisis estadístico como SPSS, nada más que éstos abrevian términos.

Tabla 10.15 Ejemplo de análisis de varianza

ANOVA Actitud hacia el entrenador del equipo de fútbol					
Fuente de variación	**Suma de cuadrados**	**Grados de libertad**	**Medias cuadráticas**	**Valor *F***	**Significancia**
Intergrupos	46 768	2	23 384	17.394	0.000
Intragrupos	793 175	590	1 344		
Total	839 943	592			

[27] Algunas pruebas son para cuando se asumen varianzas iguales y otras no, el programa indica cuáles se utilizan en cada caso.

[28] Una vez más, se elude a propósito la polémica de si la escala de Likert es de intervalos u ordinal. Si el profesor la considera ordinal, puede cambiar el ejemplo por el número de veces que han expresado públicamente su apoyo al entrenador o utilizar otro que considere pertinente en su área.

Descriptivos								
Actitud hacia el entrenador del equipo de fútbol								
					Intervalo de confianza para la media a 95%			
	N	Media	Desviación típica	Error típico	Límite inferior	Límite superior	Mínimo	Máximo
Porra Ultra	195	3.61	1.046	0.075	3.46	3.76	1	5
Porra Central	208	3.72	1.090	0.076	3.57	3.87	1	5
Porra Veteranos	190	3.07	1.331	0.097	3.88	3.26	1	5
Total	593	3.48	1.191	0.049	3.38	3.57	1	5

Comentario: la actitud de las diferentes porras hacia el entrenador es significativamente distinta, la más desfavorable es la de los veteranos (su media es de 3.07, cerrando o redondeando a décimas: 3.1).

En el ejemplo tratado en capítulos previos, cuya hipótesis es: "el consumo diario y permanente de selenio como suplemento alimenticio reduce el crecimiento de los tumores cancerígenos en mujeres que se encuentran en la etapa inicial de la enfermedad" y teniendo los tres grupos experimentales: 1) participantes a las que se les suministra un complemento alimenticio de selenio en cápsulas de 200 mg diarios, 2) participantes a las que se les administra un complemento alimenticio de selenio en cápsulas de 100 mg diarios y 3) participantes a las que no se les suministra selenio (grupo de control), se podría hacer al final del periodo experimental un *análisis de varianza* para comparar las tasas de crecimiento de los tumores entre los grupos, así como una estimación de máxima probabilidad.

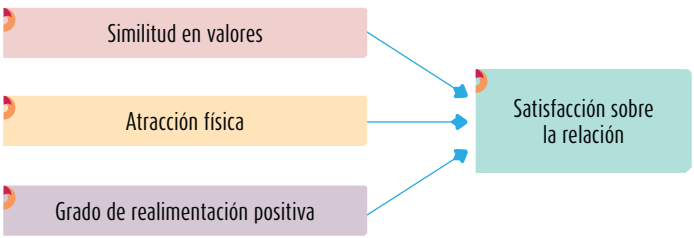
Asimismo, Lee y Guerin (2009) en su estudio para identificar si la satisfacción de la calidad del diseño ambiental del interior de áreas de trabajo u oficinas afecta significativamente la satisfacción general del espacio de trabajo por parte de sus ocupantes y su desempeño laboral, podría efectuarse un análisis de varianza por grupo de edad [30 o menos (1), 31-40 (2), 41-50 (3) y más de 50 (4)] para evaluar si difieren en cuanto a la satisfacción general sobre el espacio de trabajo.

Estadística multivariada

3 y 4 Hasta aquí hemos visto pruebas paramétricas con una sola variable independiente y una dependiente. ¿Pero qué ocurre cuando tenemos diversas variables independientes y una dependiente, varias independientes y dependientes? Se forman esquemas del tipo que se muestra en la figura 10.20.

Figura 10.20 Ejemplos de esquemas con diversas variables tanto dependientes como independientes.

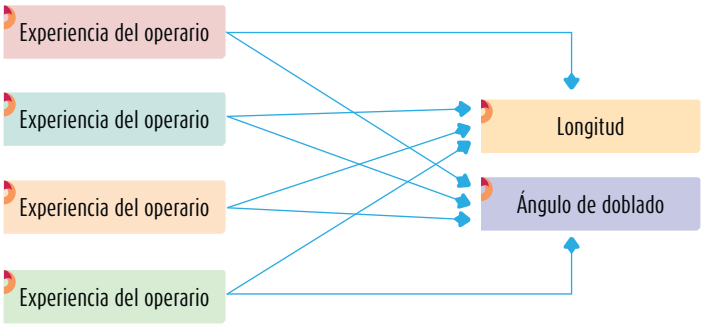
Si queremos probar la hipótesis: "la *similitud en valores*, la *atracción física* y el *grado de realimentación positiva* son factores que inciden en la *satisfacción sobre la relación* en parejas de novios cuyas edades oscilan entre los 24 y los 32 años".



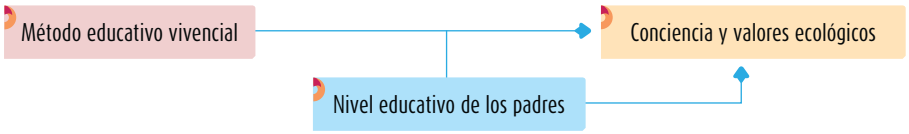
O en el estudio de Pérez, Arango y Agudelo (2009) para determinar el efecto que tienen los factores *experiencia del operario, tipo de dobladora, clase de material* utilizado y su *grosor* sobre la *longitud* y el *ángulo* de *doblado* de las piezas de metal producidas.

(*continúa*)

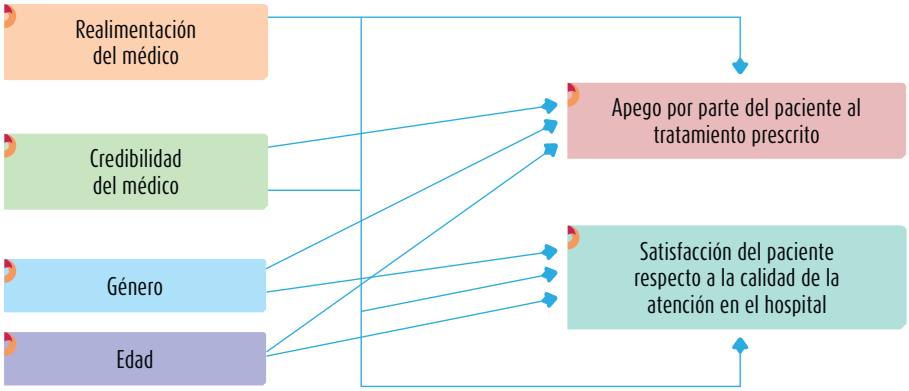
Figura 10.20 (*continuación*)



Asimismo, si pretendemos evaluar si un *método educativo* incrementa la *conciencia* y *valores ecológicos* de los estudiantes de bachillerato, controlando y analizando la influencia de la variable *nivel educativo de los padres*.



Si buscamos conocer la influencia de cuatro variables de los médicos sobre el apego al tratamiento y la satisfacción en torno a la atención por parte de sus pacientes.



Entonces, requerimos otros métodos estadísticos como los que se muestran en la tabla 10.16. Estos métodos se comentan en el capítulo 8 del centro de recursos, "Análisis estadístico: segunda parte", en análisis multivariado (Material Complementario → Capítulos → Capítulo 8).

Centro de recursos en línea

Tabla 10.16 Métodos estadísticos multivariados (se amplía en el capítulo 8 del centro de recursos en línea)

Método	Propósitos fundamentales
Análisis de varianza factorial (ANOVA de varios factores)	Evaluar el efecto de dos o más variables independientes sobre una variable dependiente.
Análisis de covarianza (ANCOVA)	Analizar la relación entre una variable dependiente y dos o más independientes, al eliminar y controlar el efecto de al menos una de estas variables independientes.
Regresión múltiple	Evaluar el efecto de dos o más variables independientes sobre una variable dependiente, así como predecir el valor de la variable dependiente con una o más variables independientes, y estimar cuál es la independiente que mejor predice las puntuaciones de la dependiente. Se trata de una extensión de la regresión lineal.

(*continúa*)

Tabla 10.16 (*continuación*)

Método	Propósitos fundamentales
Análisis multivariado de varianza (MANOVA)	Analizar la relación entre dos o más variables independientes y dos o más variables dependientes.
Análisis lineal de patrones (PATH)	Determinar y representar interrelaciones entre variables a partir de regresiones, así como analizar la magnitud de la influencia de algunas variables sobre otras, influencia directa e indirecta. Es un modelo causal.
Análisis discriminante	Construir un modelo predictivo para pronosticar el grupo de pertenencia de un caso a partir de las características observadas de cada caso (predecir la pertenencia de un caso a una de las categorías de la variable dependiente, sobre la base de dos o más independientes).
Distancias euclidianas	Evaluar la similitud entre variables (en unidades de correlación).

Análisis no paramétricos

3 y 4

Para realizar los análisis no paramétricos debe partirse de las siguientes consideraciones:[29]

1. La mayoría de estos análisis no requieren de presupuestos acerca de la forma de la distribución poblacional. Aceptan distribuciones no normales (distribuciones "libres").
2. Las variables no necesariamente tienen que estar medidas en un nivel por intervalos o de razón; pueden analizar datos nominales u ordinales. De hecho, si se quieren aplicar análisis no paramétricos a datos por intervalos o razón, éstos necesitan resumirse a categorías discretas (a unas cuantas). Las variables deben ser categóricas.

¿Cuáles son los métodos o las pruebas estadísticas no paramétricas más utilizados?

Las *pruebas no paramétricas más utilizadas* son:[30]

1. La ***chi cuadrada*** o χ^2.
2. Los coeficientes de correlación e independencia para tabulaciones cruzadas.
3. Los coeficientes de correlación por rangos ordenados de Spearman y Kendall.

Chi cuadrada Prueba estadística para evaluar hipótesis acerca de la relación entre dos variables categóricas.

¿Qué es la *Chi* cuadrada o χ^2?

Es una prueba estadística para evaluar hipótesis acerca de la relación entre dos variables categóricas.

Se simboliza: χ^2.

Hipótesis por probar: correlacionales.

Variables involucradas: dos. La prueba *Chi* cuadrada no considera relaciones causales.

Nivel de medición de las variables: nominal u ordinal (o intervalos o razón reducidos a ordinales).

Procedimiento: se calcula por medio de una *tabla de contingencia o tabulación cruzada*, que es un cuadro de dos dimensiones y cada dimensión contiene una variable. A su vez, cada variable se subdivide en dos o más categorías.

Un ejemplo de una tabla de contingencia se presenta en la tabla 10.17.

Tabla 10.17 Ejemplo de una tabla de contingencia

	Género			Total
		Masculino	Femenino	
Intención del voto	Candidata A Guadalupe Torres	40	58	98
	Candidata B Liz Almanza	32	130	162
Total		72	188	260

[29] Patrangenaru y Ellingson (2013), Chaubey (2013), Wiersma y Jurs (2008), Pett (2007), Sroka (2006) y Black (2003).
[30] Hollander, Wolfe y Chicken (2013); Howell (2011); Pett (2007), Sroka (2006); Chen y Popovich (2002); y Gibbons (1992).

Esta tabla demuestra el concepto de *tabla de contingencia* o *tabulación cruzada*. Las variables aparecen señaladas a los lados del cuadro (*intención del voto* y *género*), cada una con sus dos categorías. Se dice que se trata de una tabla 2 × 2, donde cada dígito significa una variable y el valor de éste indica el número de categorías de la variable:

2	×	2
↑		↑
Una variable con dos categorías		Otra variable con dos categorías

Un ejemplo de una tabla de contingencia 2 × 3 se muestra en la tabla 10.18. Las dos variables son: *tipo de tratamiento* contra el cáncer gástrico (tres categorías) y resultado final (dos categorías: sobrevivió o falleció). Los números que aparecen en las celdas son frecuencias. Por ejemplo: 20 pacientes que recibieron el tratamiento 1 sobrevivieron. Es necesario hacer notar que no importa cuál variable esté en la parte superior o a la izquierda, porque al final lo fundamental es que todas las categorías de una variable se crucen con todas las categorías de la otra.

Tabla 10.18 Ejemplo de una tabla de contingencia 3 × 2

		Resultado final		Total
		Sobrevivió	Falleció	
Tipo de tratamiento para el cáncer gástrico	Tratamiento 1	20	2	22
	Tratamiento 2	11	8	19
	Tratamiento 3	5	13	18
	Total	36	23	59

Otro ejemplo de relación que podrían analizarse con la *Chi* cuadrada es: *máquina utilizada* en la fabricación de tornillos (cuatro categorías: máquina 1, máquina 2, máquina 3 y máquina 4) y *calidad de la pieza* (dos categorías: defectuosa o sin defectos), para analizar diferencias por máquina.

En esencia, la *Chi* cuadrada es una *comparación* entre la *tabla de frecuencias observadas* y la denominada *tabla de frecuencias esperadas*, la cual constituye la tabla que esperaríamos encontrar si las variables fueran estadísticamente independientes o no estuvieran relacionadas (Wright, 1979). Es una prueba que parte del supuesto de "no relación entre variables" (hipótesis nula) y el investigador evalúa si en su caso esto es cierto o no, analiza si las frecuencias observadas son diferentes de lo que pudiera esperarse en caso de ausencia de correlación (Howell, 2011; Trobia, 2008; Bond, 2007b; Lane, 2006; y Platt, 2003b). La lógica es así: "si no hay relación entre las variables, debe tenerse una tabla así (la de las frecuencias esperadas). Si hay relación, la tabla que obtengamos como resultado en nuestra investigación tiene que ser muy diferente respecto de la tabla de frecuencias esperadas". Es decir, lo que se contó en comparación con lo que se esperaría del azar.

3

La *Chi* cuadrada se puede obtener a través de los programas estadísticos o mediante STATS®.

En SPSS el programa produce un resumen de los casos válidos y perdidos para cada variable (N y porcentaje) y una tabla de contingencia sencilla, como la 10.17, o bien una tabla más compleja con diversos resultados por celda (frecuencias contadas u observadas, frecuencias esperadas, porcentajes respecto a marginales y totales, etc.). Este segundo caso se muestra en el capítulo 8, "Análisis estadístico: segunda parte" (al final), en la parte sobre Chi cuadrada, que como se ha recomendado puede descargarse del centro de recursos.

El programa también proporciona el valor de *Chi* cuadrada junto con otras pruebas (pero recomendamos centrarse en el resultado de ésta), como se muestra en el siguiente ejemplo que corresponde a la tabla 10.17:

	Valor	Grados de libertad (gl)	Significancia
Chi cuadrada de Pearson	13.529	1	0.000

En este caso, el valor de *Chi* cuadrada es significativo al nivel del 0.01, es decir, se acepta la hipótesis de investigación de que existe relación entre las variables *intención del voto* y *género* (Liz Almanza gana, pero sobre todo por el voto femenino).

Algunas veces, solamente se utiliza el valor de *Chi* cuadrada simplemente para ver si hay o no relación entre las variables.

Ejemplo

Hi: "Los tres canales de televisión a nivel nacional difieren en la cantidad de programas de acción social, neutrales y antisociales que difunden". "Hay relación entre la variable *canal de televisión nacional* y la variable *emisión de programas* de acción social*, neutrales y antisociales*".

Resultados:

$$\chi^2 = 7.95$$
$$gl = 4$$

Significancia mayor de 0.05.
Conclusión: se rechaza la hipótesis de investigación y se acepta la nula. No hay relación entre las variables.

Centro de recursos en línea

El lector podrá encontrar el cálculo de la Chi cuadrada en STATS® al final del capítulo 8, "Análisis estadístico: segunda parte", del centro de recursos.

¿Qué son los coeficientes de correlación e independencia para tabulaciones cruzadas?

Además de la *Chi* cuadrada, hay *otros coeficientes* para evaluar si las variables incluidas en la tabla de contingencia o tabulación cruzada están correlacionadas. En la tabla 10.19 se describen los coeficientes más importantes para tal finalidad.[31]

Tabla 10.19 Principales coeficientes para tablas de contingencia

Coeficiente	Para cuadros de contingencia...	Nivel de medición de ambas variables	Interpretación
Phi (φ)	2×2	Nominal. Puede utilizarse con variables ordinales reducidas a dos categorías. SPSS lo muestra en cálculos para datos nominales.	En tablas 2×2 varía de 0 a 1, donde cero implica ausencia de correlación entre las variables; y uno, que hay correlación perfecta entre las variables. En tablas más grandes, *phi* puede ser mayor de 1.0, pero la interpretación es compleja. Por ello, se recomienda limitar su uso a las tablas 2×2.
Coeficiente de contingencia *C* de Pearson	Cualquier tamaño. De hecho es un ajuste de *phi* para tablas con más de dos categorías en las variables. Incluso funciona mejor con tablas de 5×5.	Nominal. Puede utilizarse con variables ordinales reducidas a dos categorías.	0 a 1, pero en tablas menores a 5×5, se acerca pero nunca alcanza el uno.
V de Cramer (*C*)	Cualquier tamaño.	Cualquier nivel de variables, pero siempre reducidas a categorías. SPSS lo muestra en cálculos para datos nominales.	0 a 1, pero el uno solamente se alcanza si ambas variables tienen el mismo número de categorías (o marginales).

(*continúa*)

[31] En SPSS encontrará otros que se incluyen por nivel de medición: Kappa, McNemar, Cochran, Riesgo, etcétera.

Tabla 10.19 (*continuación*)

Coeficiente	Para cuadros de contingencia...	Nivel de medición de ambas variables	Interpretación
Goodman-Kruskal *Lambda* o sólo *Lambda* (λ)	Cualquier tamaño.	Cualquier nivel de variables, pero siempre reducidas a categorías. SPSS lo muestra en cálculos para datos nominales.	Fluctúa entre 0 y 1, asume causalidad, lo que significa que puede predecirse a la variable dependiente definida en la tabla, sobre la base de la independiente. La versión usual de *Lambda* es asimétrica. Sin embargo, SPSS y otros programas presentan tres versiones: una simétrica y dos asimétricas (estas últimas representan a cada una de las variables considerada como dependiente). La versión simétrica es simplemente el promedio de las dos *Lambdas* asimétricas. Una prueba asimétrica presupone que el investigador puede designar cuál es la variable independiente y cuál la dependiente. En una simétrica no se asume tal causalidad.
Coeficiente de incertidumbre o entropía o *U* de Theil	Cualquier tamaño.	Cualquier nivel de variables, pero siempre reducidas a categorías. SPSS lo muestra en cálculos para datos nominales.	Fluctúa entre 0 y 1, asume causalidad, lo que significa que puede predecirse a la variable dependiente definida en la tabla, sobre la base de la independiente. Por razones históricas (de costumbre), el coeficiente se ha computado frecuentemente en términos de predecir la variable de las columnas, sobre la base de la variable de las filas.
Gamma de Goodman y Kruskal	Cualquier tamaño.	Ordinal.	Varía de −1 a +1 (−1 es una relación negativa perfecta, y +1 una relación positiva perfecta).
Tau-a, Tau-b y *Tau-c* (τa, τb, τc)	Cualquier tamaño.	Ordinal.	Varían de −1 a +1. *Tau-a* y *Tau-b* son asimétricas y *Tau-c* es simétrica.
D de Somers	Cualquier tamaño.	Ordinal.	Varía de −1 a +1.
Kappa	Cualquier tamaño.	Datos categorizados por intervalo.	Regularmente de 0 a 1.

Para profundizar en las fórmulas y premisas de estos análisis, véase Chaubey (2013), Hollander *et al.* (2013), Black (2003), Gibbons (1992) y Nie, Hull, Jenkins, Steinbrenner y Bent (1975).

¿Qué otra aplicación tienen las tablas de contingencia?

Las tablas de contingencia, además de servir para el cálculo de *Chi* cuadrada y otros coeficientes, son útiles para describir conjuntamente dos o más variables. Esto se efectúa al convertir las frecuencias observadas en frecuencias relativas o porcentajes. En una tabulación cruzada puede haber tres tipos de porcentajes respecto de cada celda.

- *Porcentaje en relación con el total de frecuencias observadas* ("*N*" *o* "*n*" *de muestra*).
- *Porcentaje en relación con el total marginal de la columna.*
- *Porcentaje en relación con el total marginal del renglón.*

Veamos con un ejemplo hipotético de una tabla 2 × 2 con las variables género y preferencia por un conductor televisivo. Las frecuencias observadas serían:

Ejemplo

		Género		
		Masculino	Femenino	
Preferencia por el conductor	*A*	25	25	50
	B	40	10	50
		65	35	100

Las celdas podrían representarse así:

a	*c*
b	*d*

Tomemos el caso de *a* (celda superior izquierda). La celda *a* (25 frecuencias observadas) con respecto al total ($N = 100$) representa 25%. En relación con el total marginal de columna (cuyo total es 65) representa 38.46% y respecto del total marginal de renglón (cuyo total es 50) significa 50%. Esto puede expresarse así:

Ejemplo

	Frecuencias observadas		
	25		
En relación con *N*	25.00%		
En relación con "*a* + *b*"	38.46%	*c*	*a* + *c* = 50
En relación con "*a* + *c*"	50.00%		
	b	*d*	*b* + *d*
	a + *b* = 65	*c* + *d*	100 = *N*

Así procedemos con cada categoría, como ocurre en la tabla 10.20.

Tabla 10.20 Ejemplo de una tabla de contingencia para describir conjuntamente dos variables

		Género		
		Masculino	Femenino	
Preferencia por el conductor	*A*	25	25	50
		25.0%	25.0%	
		38.5%	71.4%	
		50.0%	50.0%	
	B	40	10	50
		40.0%	10.0%	
		61.5%	28.6%	
		80.0%	20.0%	
		65	35	100

Algunos programas ubican los porcentajes incluidos en las celdas en otro orden. Por ejemplo, el porcentaje respecto al total lo colocan al final, pero las interpretaciones son similares.

Otros coeficientes de correlación

El coeficiente de correlación de Pearson es una estadística apropiada para variables medidas por intervalos o razón y para relaciones lineales. La *Chi* cuadrada y demás coeficientes mencionados son estadísticas adecuadas para tablas de contingencia con variables nominales, ordinales y de intervalos, pero reducidas a categorías; ahora, ¿qué ocurre si las variables de nuestro estudio son ordinales, por intervalos y de razón? O bien, una mezcla de niveles de medición, o los datos no necesariamente los disponemos en una tabla de contingencia. Existen otros coeficientes que comentaremos brevemente.

Coeficientes *rho* de Spearman y *tau* de Kendall Son medidas de correlación para variables en un nivel de medición ordinal; los individuos o unidades de la muestra pueden ordenarse por rangos.

¿Qué son los coeficientes y la correlación por rangos ordenados de Spearman y Kendall?

Los **coeficientes rho de Spearman**, simbolizado como *rs*, y **tau de Kendall**, simbolizado como *t*, son medidas de correlación para variables en un nivel de medición ordinal (ambas), de tal modo que los individuos, casos o unidades de análisis de la muestra pueden ordenarse por rangos (jerarquías). Son coeficientes utilizados para relacionar

estadísticamente escalas tipo Likert por aquellos investigadores que las consideran ordinales. Por ejemplo, supongamos que tenemos para refrescos embotellados o sodas las variables "preferencia en el sabor" y "atractivo del envase", y queremos asociarlas estadísticamente, entonces pedimos a un grupo de personas representativas del mercado que evalúen conjuntamente 10 marcas específicas y las ordenen del 1 al 10; en tanto que, "1" es la categoría o el rango máximo en ambas variables. Finalmente se obtienen los siguientes resultados en la muestra:

Marca[32]	Variable 1 Preferencia en el sabor	Variable 2 Atractivo del envase
Loy	1	2
Wiz Cola	2	5
Fan	3	1
Energizador	4	3
Maron	5	4
Manzanol	6	6
Cold	7	8
Zoda II	8	7
Frutol	9	10
Sabrosol	10	9

Para analizar tales resultados, utilizaríamos los coeficientes *rs* y *t*. Ahora bien, debe observarse que todos los refrescos o sodas tienen que jerarquizarse por rangos que contienen las propiedades de una escala ordinal (se ordena de mayor a menor). Ambos coeficientes varían de –1.0 (correlación negativa perfecta) a +1.0 (correlación positiva perfecta), considerando el 0 como ausencia de correlación entre las variables jerarquizadas. Se trata de estadísticas sumamente eficientes para datos ordinales (Howell, 2011; Khamis, 2008; Abdi, 2006; y Kraemer, 2006). La diferencia entre ellos es explicada por Nie *et al.* (1975, p. 289) de la manera siguiente: el coeficiente de Kendall (*t*) resulta un poco más significativo cuando los datos contienen un número considerable de rangos empatados. El coeficiente de Spearman *rho* parece ser una aproximación cercana al coeficiente *r* de Pearson, cuando los datos son continuos (por ejemplo, no caracterizados por un número considerable de empates en cada rango). De acuerdo con Breen y Luijkx (2010) y Creswell (2005) sirve también para analizar relaciones curvilineales.

También se interpreta su significancia igual que Pearson y otros valores estadísticos.

Otros ejemplos serían: relacionar la opinión de dos médicos respecto a la jerarquización de un grupo de pacientes en cuanto al avance de una enfermedad terminal, a fin de decidir cuáles requieren atención más urgente, o bien, dos epidemiólogos que hicieran una evaluación ordinal de las siete amenazas en salud pública para una comunidad y si hay una asociación significativa entre los dos conjuntos de filas, los funcionarios de salud se sentirán más seguros del diagnóstico de vulnerabilidad y las acciones que deben adoptar (Haug, 2007). Onwuegbuzie *et al.* (2006b) correlacionaron mediante el coeficiente *rho* el porcentaje de juegos ganados con el número de puntos anotados en una temporada para jerarquizar el desempeño de los equipos de La Liga Nacional de Fútbol Americano (NFL). Sülar y Okur (2008) relacionaron evaluaciones subjetivas hechas por 18 expertos a telas sobre los atributos de suavidad, grosor y rugosidad con mediciones objetivas de la resistencia a la tensión, flexión, corte, compresión y las propiedades de la superficie (usando diversos aparatos, entre ellos, una boquilla de pruebas), calculando, entre otras estadísticas, coeficientes de concordancia de Kendall, con el propósito de generar modelos de producción.

[32] Nombres ficticios.

¿Qué otros coeficientes hay?

Un coeficiente muy importante es el *Eta*, que es similar al coeficiente *r de Pearson*, pero con relaciones no lineales, las cuales se comentaron anteriormente. Es decir, *Eta* define la "correlación perfecta" (1.00) como curvilineal y a la "relación nula" (0.0) como la independencia estadística de las variables. Este coeficiente es asimétrico (concepto explicado en la tabla 10.19, cuando se revisa Lambda), y a diferencia de Pearson, se puede obtener un valor diferente para el coeficiente al determinar cuál variable se considera independiente y cuál dependiente. Eta^2 es interpretada como el porcentaje de la varianza en la variable dependiente explicado por la independiente. El investigador puede calcular *Eta* de las dos maneras: al cambiar la definición de la independiente y dependiente, luego promediar los dos coeficientes y obtener uno simétrico. *Eta* puede trabajarse en tablas de contingencia. Otros coeficientes se describen en la tabla 10.21.

Tabla 10.21 Otros coeficientes de correlación

Coeficiente	Nivel de medición de las variables	Ejemplos	Interpretación
Biserial (r_b)	Una ordinal y la otra por intervalos o razón.	Jerarquía en la organización y motivación. Edad y nivel de depresión	−1.00 (correlación negativa perfecta). 0.0 (ausencia de relación). +1.00 (correlación positiva perfecta).
Biserial por rangos (r_{rb})	Una variable nominal y la otra ordinal.	Escuela de procedencia (pública-privada) y rango en una prueba de un idioma extranjero (alto, medio, bajo). Género y jerarquía laboral.	−1.00 (correlación negativa perfecta). 0.0 (ausencia de relación). +1.00 (correlación positiva perfecta).
Biserial puntual (r_{pb})	Una variable por inter-valos o razón y la otra nominal.	Motivación al estudio y licenciatura (Economía, Derecho, Administración, etcétera). Número de cigarrillos fumados diariamente y desarrollo de cáncer pulmonar (presencia-ausencia de la enfermedad).	−1.00 (correlación negativa perfecta). 0.0 (ausencia de relación). +1.00 (correlación positiva perfecta).
Tetracórico o *Tetrachoric*	Las dos dicotómicas, no necesariamente expresadas en tablas. Es utilizado sobre todo cuando las variables son de intervalo o razón y han sido dicotomizadas.	Género y afiliación/no afiliación a un partido político. Decisión de abortar y creencia-no creencia en un ser supremo.	−1.00 (correlación negativa perfecta). 0.0 (ausencia de relación). +1.00 (correlación positiva perfecta).

Existen muchos más coeficientes, pero tal vez los más importantes son los señalados. Lo mejor de todo es que los programas computacionales de análisis estadístico los calculan, lo único que tenemos que hacer es interpretarlos y verbalizar sus resultados con comentarios.

Una vista general a los procedimientos o pruebas estadísticas

3 Ahora, presentamos un par de tablas finales (10.22 y 10.23) sobre los principales métodos estadísticos. En la primera se considera: *a*) *el tipo de pregunta de investigación* (descriptiva, de diferencia de grupos, correlacional o causal), *b*) *el número de variables involucradas*, *c*) *nivel de medición de las variables* o *tipo de datos* y *d*) en comparación de grupos, si son *muestras independientes* o *correlacionadas*. En este último punto, las muestras independientes se seleccionan de manera que no exista ninguna relación entre los casos de las muestras; por ejemplo, un grupo experimental y uno de control en un experimento. No hay ningún emparejamiento de las observaciones entre las muestras. Mientras que en las correlacionadas sí existe una relación entre las unidades o participantes de las muestras; por ejemplo, el mismo grupo antes y después de un tratamiento experimental, preprueba y posprueba. La

segunda, tomando en cuenta la naturaleza de la pregunta de investigación, el número de variables independientes, dependientes y de control, el tipo de variable (categórica-continua) y la distribución. Con ellas, se pretende que el alumno, junto con su profesor o profesora, seleccione las pertinentes para efectuar sus análisis de los datos.

Algunas de las pruebas o métodos estadísticos no fueron desarrollados en el capítulo y varios se encuentran en el capítulo 8 del centro de recursos en línea ("Análisis estadístico: segunda parte"), el cual podrá encontrar en: Material Complementario → Capítulos.

Tabla 10.22 Elección de los procedimientos estadísticos o pruebas[33]

1. Pregunta de investigación: Descriptiva	**Procedimiento o prueba**
• Datos nominales	Moda
• Datos ordinales	Mediana, moda
• Datos por intervalos o razón	Media, mediana, moda, desviación estándar, varianza y rango
2. Pregunta de investigación: diferencias de grupos	
a) Dos variables o grupos	
a.1. Muestras correlacionadas	
• Datos nominales	Prueba de McNemar
• Datos ordinales	Prueba de Wilcoxon para pares de rangos
• Datos por intervalos o razón	Prueba *t* para muestras correlacionadas
a.2. Muestras independientes	
• Datos nominales	*Chi* cuadrada
• Datos ordinales	Prueba Mann-Whitney U o prueba Kolmogorov-Smirnov para dos muestras
• Datos por intervalos o razón	Prueba *t* para muestras no correlacionadas o independientes
b) Más de dos variables o grupos	
b.1. Muestras correlacionadas	
• Datos nominales	Prueba Q de Cochran
• Datos ordinales	Análisis de varianza de Friedman en dos vías
• Datos por intervalos o razón	Análisis de varianza (ANOVA)
• Datos por intervalos o razón, control de efectos de otra variable independiente	Análisis de covarianza (ANCOVA)
b.2. Muestras independientes	
• Datos nominales	*Chi* cuadrada para *k* muestras independientes
• Datos nominales u ordinales (categóricos) y de intervalos-razón	*Chi* cuadrada de Friedman
• Datos ordinales	Análisis de varianza en una vía de Kruskal-Wallis (ANOVA)
• Datos por intervalos o razón	Análisis de varianza (ANOVA)
3. Pregunta de investigación: correlacional	
a) Dos variables	
• Datos nominales	Coeficiente de contingencia o *Phi*
• Datos ordinales	Coeficiente de rangos ordenados de Spearman o coeficiente de rangos ordenados de Kendall
• Datos por intervalos o razón	Coeficiente de correlación de Pearson (producto-momento)
• Una variable independiente y una dependiente (ambas de intervalos o razón)	Regresión lineal
• Datos por intervalos y nominales u ordinales	Coeficiente biserial puntual Coeficiente biserial
• Datos por intervalos y una dicotomía artificial en una escala ordinal (la dicotomía es artificial porque subyace una distribución continua)	
b) Más de dos variables	Análisis discriminante
• Datos nominales	Análisis de correlación parcial por rangos de Kendall
• Datos ordinales	Coeficiente de correlación parcial o múltiple, R^2
• Datos por intervalos o razón	

(*continúa*)

[33] Adaptado de Mertens (2005).

Tabla 10.22 (*continuación*)

4. Pregunta de investigación: causal o predictiva • Diversas independientes y una dependiente (las independientes en cualquier nivel de medición, la dependiente en nivel por intervalos o razón). Cuando las independientes son nominales u ordinales se convierten en variables "dummy" • Diversas independientes y dependientes • Agrupamiento (membresía de todos los datos) • Estructuras y redes causales	Regresión múltiple Análisis multivariado de varianza (MANOVA) Análisis discriminante (en una vía, jerárquico o factorial, de acuerdo con el número de variables involucradas) Análisis de patrones o vías (*path analysis*)
5. Pregunta de investigación: estructura de variables o validación de constructo Las variables deben estar por intervalos o razón	Análisis de factores

Tabla 10.23 Criterios para elegir las pruebas estadísticas[34]

Naturaleza de la pregunta de investigación	Número de variables independientes	Número de variables dependientes	Número de variables de control (covariables)	Tipo de variable: independiente/ dependiente	Distribución	Prueba
Comparación de grupos	1	1	0	Categórica/ continua	Normal	Prueba *t*
Comparación de grupos	1 o más	1	0	Categórica/ continua	Normal	Análisis de varianza
Comparación de grupos	1 o más	1	1	Categórica/ continua	Normal	Análisis de covarianza
Comparación de grupos	1	1	0	Categórica/ continua	No normal	Prueba de Mann-Whitney U
Relación entre grupos o categorías	1	1	0	Categórica/ categórica	No normal	*Chi* cuadrada
Correlación entre variables	La prueba no considera a una variable como independiente y la otra como dependiente, sólo establece el grado de relación. La causalidad la establece el investigador			Continua/ continua	Normal	Correlación de Pearson
Correlación entre variables	La prueba no considera a una variable como independiente y la otra como dependiente, sólo establece el grado de relación. La causalidad la establece el investigador			Categórica/ categórica	No normal	Correlación de Kendall o Spearman
Correlación entre variables	3 o más. La causalidad la establece el investigador. En la regresión múltiple, el coeficiente representa el porcentaje de variabilidad de la variable dependiente que explica el modelo de regresión			Continuas	Normal	Coeficiente de correlación múltiple (R)
Relación causal entre variables	1	1	0	Continua/ continua	Normal	Regresión lineal
Relación causal entre variables	2 o más	1	0	Continua/ continua	Normal	Regresión múltiple

Paso 6: realizar análisis adicionales

5 Este paso implica simplemente que una vez realizados nuestros análisis, es posible que decidamos ejecutar otros análisis o pruebas extras para confirmar tendencias y evaluar los datos desde diferentes ángulos. Por ejemplo, podemos en una tabla de contingencia calcular primero *Chi* cuadrada y luego *Phi*, *Lambda*, *T* de Cramer (*C*) y el coeficiente de contingencia. O después de un ANOVA, efectuar

[34] Adaptado de Creswell (2009) y Pett (2007).

www.elosopanda.com | jamespoetrodriguez.com

los contrastes posteriores que consideremos apropiados. Resulta este paso un momento clave para verificar que no se nos haya olvidado un análisis pertinente. En esta etapa regularmente se eligen los análisis multivariados.

Paso 7: preparar los resultados para presentarlos

Se recomienda, una vez que se obtengan los resultados de los análisis estadísticos (tablas, gráficas, cuadros, etc.), las siguientes actividades, sobre todo para quienes se inician en la investigación:

1. Revisar cada resultado [análisis general → análisis específico → valores resultantes (incluida la significación) → tablas, diagramas, cuadros y gráficas].
2. Organizar los resultados (primero los descriptivos, por variable del estudio; luego los resultados relativos a la confiabilidad y la validez; posteriormente los inferenciales, que se pueden ordenar por hipótesis o de acuerdo con su desarrollo).
3. Cotejar diferentes resultados: su congruencia y en caso de inconsistencia lógica volverlos a revisar. Asimismo, se debe evitar la combinación de tablas, diagramas o gráficas que repitan datos. Por lo común, columnas o filas idénticas de datos no deben aparecer en dos o más tablas. Cuando éste es el caso, debemos elegir la tabla o elemento que ilustre o refleje mejor los resultados y sea la opción que presente mayor claridad. Una buena pregunta en este momento del proceso es: ¿qué valores, tablas, diagramas, cuadros o gráficas son necesarias?, ¿cuáles explican mejor los resultados?
4. Priorizar la información más valiosa (que es en gran parte resultado de la actividad anterior), sobre todo si se van a producir reportes ejecutivos y otros más extensos.
5. Copiar o "formatear" las tablas en el programa con el cual se elaborará el reporte de la investigación (procesador de textos —como Word— o uno para presentaciones, como Power Point, Flash, Prezi). Algunos programas como SPSS y Minitab permiten que se transfieran los resultados (tablas, por ejemplo) directamente a otro programa (copiar y pegar). Por ello, resulta conveniente usar una versión del programa de análisis que esté en el mismo idioma que se empleará para escribir el reporte o elaborar la presentación. Aunque, de no ser así, el texto de las tablas y gráficas puede modificarse, únicamente es más tardado.
6. Comentar o describir brevemente la esencia de los análisis, valores, tablas, diagramas, gráficas.
7. Volver a revisar los resultados.
8. Y, finalmente, elaborar el reporte de investigación.

En el centro de recursos se encuentran más ejemplos de estudios con diferentes análisis tratados en este capítulo y en el capítulo 8 adicional de la página, "Análisis estadístico: segunda parte". Se incluye al final del presente capítulo una secuencia de análisis en Minitab con la investigación de la televisión y el niño, y en el capítulo 8 adicional una secuencia de análisis en SPSS con un estudio del clima organizacional.

SPSS®
Centro de recursos en línea

Resumen

- El análisis cuantitativo de los datos se efectúa mediante la matriz de datos, la cual está guardada como archivo.
- Los pasos más importantes en el análisis de los datos son:
  - Decidir el programa de análisis de los datos por utilizar.
  - Explorar los datos obtenidos en la recolección:
    - *a*) Analizar descriptivamente los datos por variable del estudio.
    - *b*) Visualizar los datos por variable.
  - Evaluar la confiabilidad y validez del instrumento o instrumentos de medición utilizados.
  - Analizar e interpretar mediante pruebas estadísticas las hipótesis planteadas (análisis estadístico inferencial).
  - Realizar análisis adicionales.
  - Preparar los resultados para presentarlos.
- Los análisis estadísticos se llevan a cabo mediante programas computacionales, los más conocidos son: IBM SPSS, Minitab y SAS.
- El tipo de análisis o pruebas estadísticas depende del nivel de medición de las variables, las hipótesis y el interés del investigador.

capítulo

11

El reporte de resultados del proceso cuantitativo

No existe la investigación perfecta, pero debemos tratar de demostrar que hicimos nuestro mejor esfuerzo. El reporte de investigación es la oportunidad para ello.

Roberto Hernández-Sampieri

Proceso de investigación cuantitativa

Paso 10 Elaborar el reporte o informe de resultados

- Definición del usuario.
- Selección del tipo de reporte a presentar: formato y contexto académico o no académico, dependiendo del usuario.
- Elaboración del reporte siguiendo un estilo de publicaciones y del material adicional correspondiente.
- Presentación del reporte.

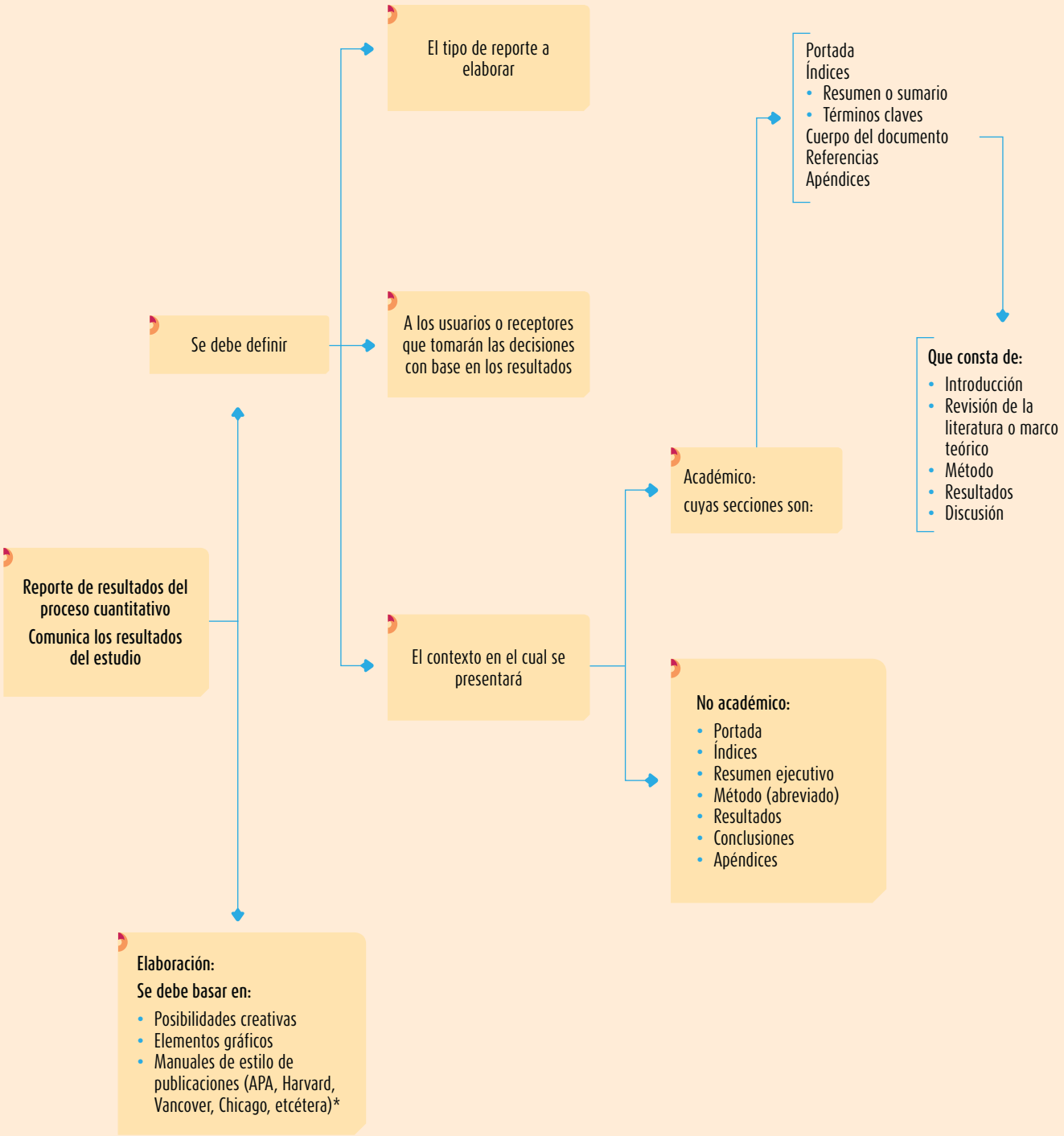
Al terminar este capítulo, el alumno será capaz de:

1. Entender el papel tan importante que cumple el usuario de la investigación en la elaboración del reporte o informe de resultados.
2. Reconocer los tipos de informes de resultados en la investigación cuantitativa.
3. Comprender los elementos que integran un reporte de investigación cuantitativa.

Síntesis

En el capítulo se comenta la importancia que tienen los usuarios en la presentación de resultados. Éstos son quienes toman decisiones con base en los resultados e inferencias de la investigación; por ello, la presentación debe adaptarse a sus características y necesidades.

Se mencionan dos tipos de reportes: académicos y no académicos, así como los elementos o secciones más comunes que integran un reporte producto de la investigación cuantitativa.



Nota: En el centro de recursos en línea puede descargar y consultar el capítulo 11 adicional, que contiene un resumen de cómo incluir la bibliografía básica según varios estilos de publicación y sus páginas de internet. Asimismo, pueden descargarse dos manuales sencillos, uno sobre el estilo APA y otro sobre el estilo Vancouver, y el programa SISI para generar, incluir y organizar referencias bibliográficas, tanto en el texto —citas— como al final en el listado o bibliografía —referencias—, basadas en el estilo APA.

Centro de recursos en línea

Antes de elaborar el reporte de investigación, se definen los receptores o usuarios y el contexto

Se ha llevado a cabo una investigación cuantitativa y se generaron resultados que se encuentran en tablas, gráficas, cuadros, diagramas, etc.; pero el proceso aún no termina. Es necesario comunicar los resultados mediante un reporte, el cual puede adquirir diferentes formatos: un libro o un capítulo, un artículo para una revista académica, un diario de divulgación general, una presentación en computadora, un documento técnico, una tesis o disertación, un DVD, etc. En cualquier caso, debemos describir la investigación realizada y los descubrimientos producidos.

Lo primero entonces es definir el tipo de reporte que es necesario elaborar, esto depende de varias precisiones:

1. las razones por las cuales surgió la investigación,
2. los **usuarios** del estudio y
3. el contexto en el cual se habrá de presentar.

Usuarios Personas que toman decisiones con base en los resultados de la investigación; por ello, la presentación debe adaptarse a sus necesidades.

Por tanto, es necesario que antes de comenzar a preparar el reporte, el investigador reflexione respecto de las siguientes preguntas: ¿cuál fue el motivo o los motivos que originaron el estudio? ¿Cuál es el contexto en que habrán de presentarse los resultados? ¿Quiénes son los usuarios de los resultados? ¿Cuáles son las características de tales usuarios? La manera en que se presenten los resultados dependerá de las respuestas a dichas preguntas.

1

Si el motivo fue elaborar una tesis para obtener un grado académico, el panorama es claro: el formato del reporte debe ser, justamente, una tesis de acuerdo con el grado que se cursó (licenciatura o pregrado, maestría o doctorado) y los lineamientos son los establecidos por la institución educativa donde se habrá de presentar, el contexto será académico y los usuarios serán, en primera instancia, los sinodales o miembros de un jurado y, posteriormente, otros alumnos y profesores de la propia universidad y otras organizaciones educativas. Si se trata de un trabajo solicitado por un profesor para una materia o curso, el formato es un informe académico cuyo usuario principal es el maestro que lo encargó y los beneficiarios o lectores inmediatos son los compañeros que cursan la misma asignatura, para que después se agreguen como usuarios otros estudiantes de la escuela o facultad de nuestra institución y de otras universidades. En caso de que la razón que originó el estudio fue la solicitud de una empresa para que se analizara determinado aspecto que interesa a sus directivos. El reporte será en un contexto no académico y los usuarios básicamente son un grupo de ejecutivos de la organización en cuestión que utilizará los datos para tomar ciertas decisiones.

También en ocasiones la investigación obedece a varios motivos y tiene diferentes interesados (imaginemos que realizamos un estudio pensando en diversos productos y usuarios: un artículo que se someta a consideración para ser publicado en una revista científica, una ponencia para ser presentada en un congreso, un libro, etc.). En este caso, suele primero elaborarse un documento central para después desprender de éste distintos subproductos.

Vamos primero a considerar a los usuarios de la investigación, los contextos en que puede presentarse, los estándares que regularmente se contemplan al elaborar un reporte y que debemos tomar en cuenta, así como el tipo de informe que se utiliza en cada caso. Todos estos elementos se resumen en la tabla 11.1.

Tabla 11.1 Usuarios, contextos y estándares para la investigación[1]

Usuarios	Contextos comunes posibles	Estándares que normalmente aplican para elaborar el reporte	Tipo de reporte o informe
Académicos de la propia institución educativa: profesores, asesores, miembros de comités y jurados, alumnos (tesis y disertaciones, estudios institucionales para sus propias publicaciones o de interés para la comunidad universitaria).	Académico	• Lineamientos utilizados en el pasado para regular las investigaciones en la escuela o a nivel institucional. Es común que haya un manual de la institución. • Lineamientos individuales de los decanos y profesores-investigadores de la escuela, facultad o departamento.	• Tesis y disertaciones. • Informes de investigación. • Presentaciones audiovisuales (Power Point, Flash, Prezi, Dreamweaver, SlideRocket, etcétera). • Libro o capítulo de libro. • Artículo para revista académica interna.

(*continúa*)

[1] Adaptado de Creswell (2005, p. 258).

Tabla 11.1 (*continuación*)

Usuarios	Contextos comunes posibles	Estándares que normalmente aplican para elaborar el reporte	Tipo de reporte o informe
Editores y revisores de revistas científicas (*journals*) y comunidades académicas lectoras.	Académico	• Lineamientos publicados por el editor o comité editorial de la revista. En ocasiones se diferencian por su enfoque: si son investigaciones cuantitativas, cualitativas o mixtas. Es común que se denominen "normas o instrucciones para los autores", "someter un manuscrito", "guía para autores", etc. El planteamiento de nuestro estudio debe encuadrar dentro del tema o temas de la revista y a veces en el volumen en cuestión (que puede ser anual o bianual) o el número (mensual, bimensual, trimestral, cuatrimestral o anual).	• Artículos que reportan resultados de una investigación (largos o cortos). • Informes de evaluaciones o diagnósticos. • Reportes de metaanálisis. • Artículos conceptuales o teoréticos. • Ensayos. • Informes de estudios de caso. • Opiniones fundamentadas. • Revisiones de libros.
Revisores de ponencias y papeles para congresos académicos (ponencias y presentaciones en seminarios, foros en internet, páginas web, premios a la investigación, etcétera).	Académico	• Lineamientos o estándares definidos en la convocatoria del congreso, foro o certamen. Estos estándares son para el escrito que se presenta o publica, así como para los materiales adicionales requeridos (por ejemplo, presentación audiovisual, video, resumen gráfico para cartel). El planteamiento de nuestro estudio debe encuadrar dentro del tema del congreso y tenemos que ajustarnos a la normatividad definida para las ponencias.	• Ponencias escritas o papeles. • Póster o cartel. • Presentación en un archivo para PC, laptop, tableta o similar.

Elaboradores de políticas, ejecutivos o funcionarios que toman decisiones (empresas, organizaciones gubernamentales y organizaciones no gubernamentales).	Académico
	No académico (regularmente el caso de las empresas)
		• Lineamientos lógicos o estándares utilitarios:
		  • Informe breve, cuyos resultados sean fáciles de entender.
		  • Orientación más bien visual del contenido (gráficas, cuadros, etc.; solamente los elementos más importantes).
		  • Posibilidad de aplicar los resultados de manera inmediata.
		  • Claridad de ideas.
			• Resumen ejecutivo.
			• Informe técnico.
			• Presentaciones
			• audiovisuales.

Profesionales y practicantes dentro del campo donde se inserta el estudio.	Académico
	No académico (regularmente el caso de las empresas)
		• Lineamientos lógicos o estándares pragmáticos:
		  • Relevancia del problema estudiado.
		  • Orientación más bien visual del contenido (gráficas, cuadros, etc.; sólo los elementos más importantes).
		  • Resultados fácilmente identificables y aplicables.
		  • Sugerencias prácticas y concretas para implementar.
			• Resumen ejecutivo.
			• Informe técnico.
			• Presentaciones
			• audiovisuales.

Opinión pública no especializada (estudiantes de primeros ciclos, padres de familia, grupos de la sociedad en general).	No académico	• Estándares centrados en la sencillez de los resultados, su importancia para un grupo de la sociedad o ésta en su conjunto:
		  • Brevedad.
		  • Claridad.
		  • Aplicabilidad a situaciones cotidianas.
		  • Orientación más bien visual del contenido (gráficas, cuadros, etc.; pocos elementos, dos o tres muy sencillos).
			• Artículo periodístico.
			• Libro de divulgación popular.

Los estándares son las bases para elaborar el reporte. La regulación en el campo académico casi siempre es mayor que en contextos no académicos, en los cuales no hay tantas reglas generales.

El primer aspecto que hay que considerar es la extensión. En el capítulo tres hacíamos una analogía entre la investigación y la actividad artística de pintar. El reporte equivale al lienzo. Hay usuarios

y normatividades que permiten "lienzos" más extensos que otros. Por ello, los informes varían en tamaño, pues éste depende del estudio en sí y las normas editoriales. Aunque la tendencia actual es incluir sólo los elementos y contenidos realmente necesarios.

Algunos autores, como Creswell (2005), señalan que en tesis de licenciatura y maestría los límites habituales son de 50 a 125 páginas de contenido esencial (sin contar apéndices). Las disertaciones doctorales, entre 100 y 300 páginas, y los informes ejecutivos de tres a 10 páginas. En el caso de los artículos para revistas científicas, se presentan variaciones dependiendo del área del conocimiento y el tipo de documento (Dahlberg, Wittink y Gallo, 2010). Por tanto, se revisaron las instrucciones para elaborar artículos de 300 revistas internacionales y 50 iberoamericanas.[2] En la tabla 11.2 se presentan las extensiones mínimas y máximas permitidas (en palabras y por tipo de informe, y en algunos casos en número de páginas).

Tabla 11.2 Extensiones mínimas y máximas permitidas para escritos de diversas clases en revistas científicas por áreas generales de conocimiento[3]

Tipo de reporte	*Ciencias sociales y administrativas*	*Ciencias de la salud*	*Ciencias biológicas*	*Humanidades (incluyendo Derecho)*	*Ingenierías y matemáticas*	*Arquitectura y medio ambiente*
Artículos que reportan resultados	3 000 a 10 000 palabras o 25 páginas, excepto metodología: 30 páginas y educación: 35 páginas, incluyendo todo.	2 500 a 8 000 palabras o 20-25 páginas.	5 000 palabras.	7 000 a 10 000 palabras.	5 000 a 8 000 palabras.	1 500 a 8 000 palabras o 6 a 25 páginas.
Artículos conceptuales (revisiones)	3 500 a 8 000 palabras.	3 000 a 8 000 palabras.	3 000 a 5 000 palabras.	5 000 palabras.	6 000 a 7 000 palabras.	7 000 palabras.
Estudios de caso	600 a 5 000 palabras.	500 a 1 500 palabras.	2 000 a 3 000 palabras.	2 000 a 5 000 palabras.	1 000 a 5 000 palabras.	1 500 a 7 000 palabras.
Ensayos o tutoriales	15 páginas.	1 500 palabras.	1 000 palabras.	1 000 palabras.	1 000 a 4 000 palabras.	–

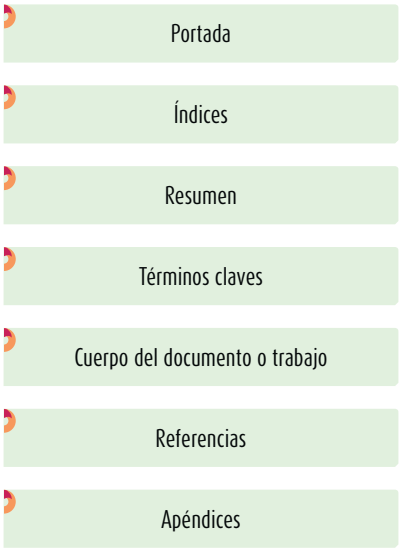
Asimismo, los informes de proyectos en desarrollo: entre 1 000 y 1 500 palabras. Los posters o carteles normalmente son de una o dos páginas de acuerdo con el tamaño que sea requerido por los organizadores del congreso. Los escritos para presentarse como ponencias suelen no exceder de 30 minutos (será necesario calcular el equivalente en páginas de acuerdo con el ritmo del orador), pero también depende del comité que organiza cada acto académico. Los artículos periodísticos regularmente no ocupan más de una página del diario, en el caso más extenso.

¿Qué apartados o secciones contiene un reporte de investigación o un reporte de resultados en un contexto académico?

Las secciones más comunes de los reportes de investigación, en la mayoría de los casos, son los que a continuación se comentan:[4]

[2] Diversas revistas iberoamericanas no tienen actualizadas sus páginas web, otras no son arbitradas, tienen normas imprecisas o no las incluyen.
[3] Los mínimos no incluyen referencias y apéndices; los máximos, sí.
[4] Creswell (2013a); *Journal of Chemical Technology and Biotechnology* (2013); *Multiple Sclerosis Journal* (2013); *Diabetes & Vascular Disease Research* (2013); Babbie (2012); American Psychological Association (2011); Dahlberg, Wittink y Gallo (2010); Mertens (2010); The Psychology Research Handbook (2006); y Williams, Unrau y Grinnell (2005). Además, el lector puede constatar que la mayoría de los journals científicos solicita esta estructura.

www.elosopanda.com | jamespoetrodriguez.com



1. Portada

Incluye el título de la investigación; el nombre del autor o los autores y de su institución o el nombre de la organización que patrocina el estudio, así como la fecha y el lugar en que se presenta el reporte. En el caso de tesis y disertaciones, las portadas varían de acuerdo con los lineamientos establecidos por la autoridad pública o la institución de educación superior correspondiente.

Se aconseja que el título del informe resulte breve pero informativo y que se eviten términos "sensacionalistas" (por ejemplo, el título: "Consecuencias psicológicas de la violación en jóvenes adolescentes" es crudo, sensacionalista e indica una nula sensibilidad ante una situación extremamente delicada. Asimismo, el título no debe ser muy corto ni muy largo. Por ejemplo, los siguientes títulos son breves pero imprecisos: "Estudio del embarazo ectópico", "El clima organizacional en las empresas peruanas", "Las preferencias de los consumidores de Montevideo". Y el siguiente título es demasiado extenso: "Análisis de la autonomía laboral, la satisfacción en el trabajo, la motivación intrínseca, el compromiso con la empresa, el involucramiento en el trabajo, la cooperación dentro de los departamentos y otros elementos del clima organizacional en cuatro empresas del giro de la construcción de tamaño mediano ubicadas en el municipio de León, Guanajuato, y su relación con la productividad, la innovación y el trabajo en equipo". Por lo regular, en el título no se incluyen nombres de instituciones ni el número de casos de estudio, pero la hipótesis central sí puede ser parte (Betkerur, 2008). Tampoco se aceptan abreviaturas (United European Gastroenterology Journal, 2013).

En cuanto al tamaño del título, varía entre 12 y 20 términos (60 a 80 caracteres).[5]

2. Índices

Normalmente, los índices son varios: primero el de contenido, que incluye capítulos, apartados y subapartados (diferenciados por numeración progresiva o tamaño y características de la tipografía). Posteriormente el índice de tablas y el índice de figuras.[6] Desde luego, los índices solamente se incluyen en reportes largos como las tesis e informes.

3. Resumen o sumario

Resumen Síntesis breve de los contenidos del estudio que permite que los lectores conozcan las generalidades de la investigación.

Constituye el contenido esencial del reporte de investigación y usualmente incluye: *a*) el planteamiento del problema (expresado en una o dos oraciones, frecuentemente

[5] Journal of Chemical Technology and Biotechnology (2013), Applied Psychological Measurement (2013), Business Communication Quarterly (2013) y Study Guides and Strategies Website (2011).
[6] Las figuras incluyen diagramas, fotografías, dibujos, sistemas de ecuaciones, esquemas y gráficas de resultados, como histogramas y diagramas de dispersión.

como objetivo o propósito), *b*) método (unidades de análisis, diseño, muestra, instrumento), *c*) resultados o descubrimientos más importantes (dos o tres, uno de éstos, la prueba de hipótesis) y *d*) las principales conclusiones e implicaciones. Debe ser comprensible, sencillo, informativo, preciso, completo, conciso y específico.

En el caso de artículos las normas de las revistas científicas solicitan sumarios de entre 100 y 300 palabras.[7] La American Psychological Association (2011) sugiere no exceder de 120. Tratándose de artículos conceptuales, se encuentran entre 100 y 150 palabras. En tesis y disertaciones, se sugiere que no exceda las 320 palabras (el estándar es de 300). Para reportes técnicos se sugiere un mínimo de 200 palabras y un máximo de 350.[8] Casi en todas las revistas académicas y tesis se exige que el resumen esté en el idioma original en que se produjo el estudio (en nuestro caso en español) y en inglés. A continuación se presenta un ejemplo.

Ejemplo

A cross-cultural examination of interpersonal communication motives in México and the United States[9]

Rebecca B. Rubin
Kent State University
Carlos Fernández Collado
Universidad Anáhuac
Roberto Hernández-Sampieri
Universidad Anáhuac

Resumen

Este estudio examina las diferencias culturales en los motivos para comunicarse de manera interpersonal, comparando una muestra de estudiantes universitarios estadounidenses con otra de mexicanos. La investigación previa indica que hay seis motivos principales para iniciar conversaciones con los demás: *placer*, *escape*, *relajación*, *inclusión*, *afecto interpersonal* y *control*. Las cuatro dimensiones de la cultura nacional reportadas por Hofstede (1980): distancia al poder, evitación de la incertidumbre, individualismo y masculinidad fueron utilizadas para predecir diferencias interculturales en tales motivos interpersonales.

Se sometieron a prueba ocho hipótesis. Los resultados indican que las puntuaciones de los universitarios mexicanos no fueron significativamente mayores que las puntuaciones de los alumnos estadounidenses en los motivos de *control*, *relajación* y *escape*, tal como se había predicho; pero sí fueron significativamente menores en cuanto a los motivos de *afecto interpersonal*, *placer* e *inclusión*. Asimismo, se presentaron correlaciones negativas significativas entre los motivos interpersonales y la edad en los datos de Estados Unidos, pero no en los datos de México. También se descubrieron correlaciones positivas significativas entre el género y los motivos de *afecto* e *inclusión*, y correlaciones negativas entre el género y el motivo de *control*, pero solamente en los universitarios estadounidenses.

La muestra de mexicanos incluyó a 225 individuos y la de estadounidenses a 504.

4. Términos claves (*keywords*)

Son términos que identifican al tipo de investigación o trabajo realizado y son útiles para ayudar a los indexadores y motores de búsqueda a encontrar los reportes o documentos pertinentes. Por lo común se solicitan entre tres y cinco términos, aunque algunas revistas permiten hasta ocho, como *Acta Sociologica*. Hay revistas iberoamericanas que aceptan hasta 10, lo cual nos parece extremo. Un término clave puede abarcar más de una palabra, por ejemplo: modelamiento latente de Markov, ecuaciones estructurales, actitudes hacia las matemáticas, medicina basada en evidencia, auditoría en comunicación, fibra de madera, etcétera.

[7] No hay relación entre el campo o disciplina y el número de palabras permitidas. Por ejemplo, de las 350 revistas analizadas, *Journal of Fire Sciences*, *European Law Journal* y *Business Communication Quarterly* solicitan como máximo entre 100 y 150 y el *Journal of Cellular Plastics* 300.
[8] Williams, Unrau y Grinnell (2005).
[9] Adaptado de Rubin, Fernández y Hernández-Sampieri (1992).

www.elosopanda.com | jamespoetrodriguez.com

5. Cuerpo del documento

Respecto a las partes que integran el cuerpo del documento, existen variantes entre los diferentes campos del conocimiento, asociaciones científicas, normas editoriales y perspectivas del investigador. Desde luego, por cuestiones de espacio no sería viable tratar todas estas variantes, nos limitaremos a mencionar la principal y después comentaremos cada parte o sección, citando algunas diferencias.

- *Introducción*: abarca los antecedentes (tratados de manera breve, concreta y específica), el planteamiento del problema (objetivos y preguntas de investigación, así como la justificación del estudio), un sumario de la revisión de la literatura, el contexto de la investigación (cómo, cuándo y dónde se realizó), las variables y los términos de la investigación, lo mismo que las limitaciones de ésta. Es importante que se comente la utilidad del estudio para el campo académico y profesional. Creswell (2005) lo denomina el planteamiento del problema y agrega las hipótesis.

  Diversas revistas académicas en todos los campos, asociaciones científicas e instituciones de educación superior incluyen en la introducción el marco teórico o revisión de la literatura, particularmente en artículos de investigación empírica.

  Laflen (2001) recomienda una serie de preguntas para elaborar la introducción: ¿qué descubrió o probó la investigación? ¿En qué clase de problema se trabajó, cómo se trabajó y por qué se trabajó de cierta manera?, ¿qué motivó el estudio? ¿Por qué se escribe el reporte? ¿Qué debe saber o entender el lector al terminar de leer el reporte?
- *Revisión de la literatura* (*marco teórico*): en ésta se incluyen y comentan las teorías que se manejaron y los estudios previos que fueron relacionados con el planteamiento, se hace un sumario de los temas y hallazgos más importantes en el pasado y se señala cómo nuestra investigación amplía la literatura actual. Finalmente, tal revisión nos debe responder la pregunta: ¿dónde estamos ubicados actualmente en cuanto al conocimiento referente a nuestras preguntas y objetivos?
- *Método*: esta parte del reporte describe cómo fue llevada a cabo la investigación, e incluye:

  –Enfoque (cuantitativo, cualitativo o mixto).
  –Contexto de la investigación (lugar o sitio y tiempo, así como accesos y permisos).
  –Casos, universo y muestra (tipo, procedencia, edades, género o aquellas características que sean relevantes de los casos; descripción del universo y la muestra, y procedimiento de selección de la muestra).
  –Diseño utilizado (experimental o no experimental —diseño específico—, así como intervenciones, si es que se utilizaron).
  –Procedimiento(s) (un resumen de cada paso en el desarrollo de la investigación). Por ejemplo, en un experimento se describe la manera de asignar los casos a los grupos, las instrucciones, los materiales, las manipulaciones experimentales y cómo transcurrió el experimento. En una encuesta se refiere cómo se contactó a los participantes y se realizaron las entrevistas. En este rubro se incluyen los problemas enfrentados y la forma en que se resolvieron.
  –Descripción detallada de los procesos de recolección de los datos y qué se hizo con los datos una vez obtenidos.
  –En cuanto a la recolección, es necesario describir qué datos fueron recabados, cuándo fueron recogidos y cómo: forma de recolección o instrumentos de medición utilizados, con reporte de la confiabilidad, validez y objetividad, así como las variables o conceptos, eventos, situaciones y categorías.

Cabe señalar que en áreas de conocimiento como las ingenierías, las ciencias químicas y biológicas, las ciencias de la salud y la arquitectura, suele agregarse un apartado en esta sección de método, denominado: "Materiales", que incluye equipos, utensilios, aparatos, software y otras tecnologías utilizadas, de las cuales debe precisarse su uso e influencia en el estudio. Incluso, algunas revistas lo consideran una parte del cuerpo del documento (por ejemplo, *Journal of Dental Research*, *Journal of Chemical Technology and Biotechnology* y *Multiple Sclerosis Journal*).

- *Resultados*: son producto del análisis de los datos. Compendian el tratamiento estadístico que se dio a los datos. Regularmente el orden es: *a*) análisis descriptivos de los datos, *b*) análisis inferen-

ciales para responder a las preguntas o probar hipótesis (en el mismo orden en que fueron formuladas las hipótesis o las variables). La American Psychological Association (2011) recomienda que primero se describa de manera breve la idea principal que resume los resultados o descubrimientos, y posteriormente se reporten con detalle los resultados. Es importante destacar que en este apartado no se incluyen conclusiones ni sugerencias, así como tampoco se explican las implicaciones de la investigación. Esto se hace en el siguiente apartado.

En la sección de resultados el investigador se limita a describir sus hallazgos. Una manera útil de hacerlo es mediante tablas, cuadros, gráficas, dibujos, diagramas, mapas y figuras generadas por el análisis. Son elementos que sirven para organizar los datos, de tal manera que el usuario o lector los pueda leer y decir: "me queda claro que esto se vincula con aquello, con esta variable ocurre tal cuestión." Cada uno de dichos elementos debe ir numerado en arábigo o romano (por ejemplo: cuadro 1, cuadro 2... cuadro *k*; gráfica o diagrama 1, gráfica o diagrama 2... gráfica o diagrama *k*, etc.) y con el título que la identifica. Wiersma y Jurs (2008) recomiendan los siguientes puntos para elaborar tablas estadísticas:

a) El *título* debe especificar el contenido de la tabla, así como tener un *encabezado* y los *subencabezados* necesarios (por ejemplo, columnas y renglones, diagonales, etcétera).
b) No debe mezclarse una cantidad inmanejable de estadísticas (por ejemplo, incluir medias, desviaciones estándar, correlaciones, razón *F*, etc., en una misma tabla).
c) En cada tabla se deben *espaciar los números y las estadísticas incluidas* (tienen que ser legibles).
d) En tesis e informes y de ser posible, habrá que *limitar cada tabla a una sola página.*
e) Los formatos de las tablas tienen que ser coherentes y homogéneos dentro del reporte (por ejemplo, no incluir en una tabla cruzada las categorías de la variable dependiente en columnas y en otra tabla colocar las categorías de la variable dependiente en renglones).
f) Las *categorías de las variables deben distinguirse* claramente entre sí.

La mejor regla para elaborar una tabla es organizarla lógicamente y eliminar la información que pueda confundir al lector. Al incluir pruebas de significancia: *F*, *Chi cuadrada, r*, etc., debe incorporarse información respecto de la magnitud o el valor obtenido de la prueba, los grados de libertad, el nivel de confianza (*alfa* = α) y la dirección del efecto (American Psychological Association, 2011). Asimismo, tendrá que especificarse si se acepta o se rechaza la hipótesis de investigación o la nula en cada caso.

Centro de recursos en línea

Recomendamos a los lectores consultar los ejemplos de investigación incluidos en el centro de recursos de esta obra y revisar la forma como se presentan las tablas.

Cuando los *usuarios*, receptores o lectores son personas con conocimientos sobre estadística no es necesario explicar en qué consiste cada prueba, sólo habrá que mencionarla y comentar sus resultados (que es lo normal en ambientes académicos). Si el usuario carece de tales conocimientos, no tiene caso incluir las pruebas estadísticas, a menos que se expliquen con suma sencillez y se presenten los resultados más comprensibles. En este caso, las tablas se describen.

En el caso de diagramas, figuras, mapas cognoscitivos, esquemas, matrices y otros elementos gráficos, también debe seguirse una secuencia de numeración y observar el principio básico: *una buena figura es sencilla, clara y no estorba la continuidad de la lectura*. Las tablas, los cuadros, las figuras y los gráficos tendrán que enriquecer el texto; en lugar de duplicarlo, comunican los hechos esenciales, son fáciles de leer y comprender, a la vez que son coherentes.

La mayoría de revistas requieren las figuras y tablas como archivos *jpeg*, *tiff* o *eps* (resolución mínima de 300 dpi).

- *Discusión* (conclusiones, recomendaciones, limitaciones e implicaciones)*:* en esta parte se: *a*) derivan conclusiones, *b*) explicitan recomendaciones para otros estudios (por ejemplo, sugerir nuevas preguntas, muestras, instrumentos, líneas de investigación, etc.) y se indica lo que sigue y lo que debe hacerse, *c*) generalizan los resultados a la población, *d*) evalúan las implicaciones del estudio, *e*) establece la manera como se respondieron las preguntas de investigación, así como si se cumplieron o no los objetivos, *f*) relacionan los resultados con los estudios existentes (vincular con el marco teórico y señalar si nuestros resultados coinciden o no con la literatura previa, en qué sí y

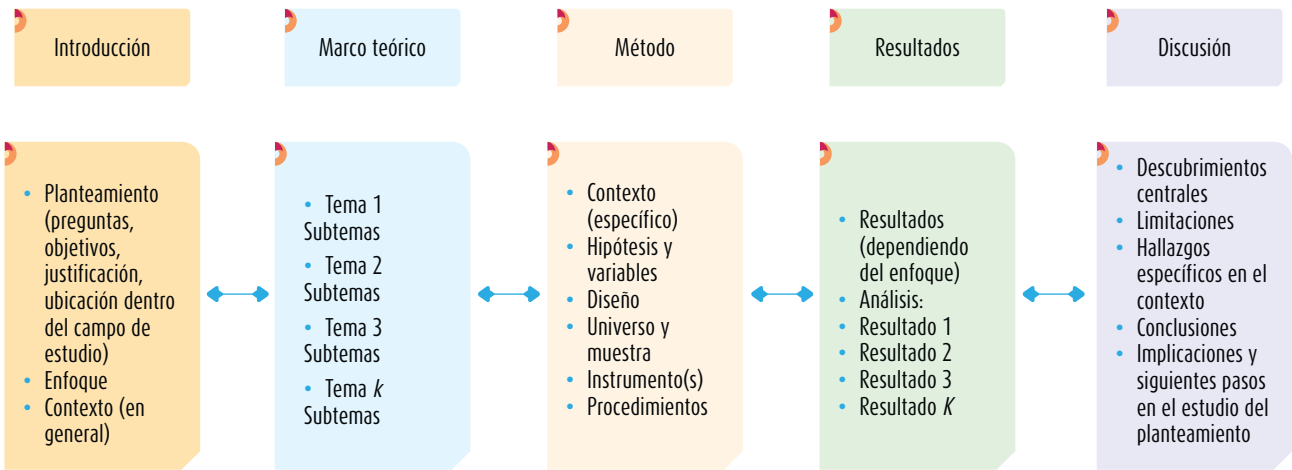
www.elosopanda.com | jamespoetrodriguez.com

en qué no), *g*) reconocen las limitaciones de la investigación (en el diseño, muestra, funcionamiento del instrumento, alguna deficiencia, etc., con un alto sentido de honestidad y responsabilidad), *h*) destaca la importancia y significado de todo el estudio y la forma como encaja en el conocimiento disponible, *i*) explican los resultados inesperados y *j*) cuando no se probaron las hipótesis es necesario señalar o al menos especular sobre las razones. Al elaborar las conclusiones es aconsejable verificar que estén los puntos necesarios aquí vertidos. Y recordar que **no** se trata de repetir los resultados, sino de resumir los más importantes y su significado. Desde luego, las conclusiones deben ser congruentes con los datos. La adecuación de éstas respecto de la generalización de los resultados deberá evaluarse en términos de aplicabilidad a diferentes muestras y poblaciones. Si el planteamiento cambió, es necesario explicar por qué y cómo se modificó. Esta parte debe redactarse de tal manera que se facilite la toma de decisiones respecto de una teoría, un curso de acción o una problemática. El reporte de un experimento tiene que explicar con claridad las influencias de los tratamientos.

Congruencia entre partes del cuerpo de documento o apartados

Un aspecto que cabe destacar de todo informe es que debe haber una elevada congruencia entre las diferentes partes que integran el documento. Por ello, al elaborar el reporte, aunque nos concentremos en la redacción de un apartado, es indispensable tener en mente el resto de las secciones y asegurar que haya vinculación entre éstas. Un informe de investigación puede concebirse como un esquema, tal como se presenta en la figura 11.1.

Figura 11.1 Esquema de bosquejo general para un reporte de investigación.



La congruencia implica, por ejemplo, que todos los temas y subtemas del marco teórico se encuentren estrechamente relacionados con el planteamiento y hayan sido incluidos en los resultados, que los descubrimientos, hallazgos y conclusiones estén asociados con los análisis y resultados y, desde luego, con los apartados del marco teórico, que todos los elementos de la discusión se refieran al planteamiento (objetivos, preguntas y justificación), que las hipótesis y variables se localicen en el marco teórico (se presenten estudios sobre ellas) y se discutan en las conclusiones, etc. Las variables del planteamiento deben estar presentes en todos los temas del marco teórico y en las hipótesis, tienen que ser las medidas por el instrumento o los instrumentos, deben estar descritas y relacionadas en los resultados e implicadas en la discusión.

Las partes o secciones del cuerpo del reporte tienen que ser como "espejos que deben reflejar una misma imagen" (Dahlberg, Wittink y Gallo, 2010). Esta labor nos ayuda a mantener un reporte congruente de principio a fin. Todo conectado a la literatura sobre el planteamiento del problema en cuestión.

Se recomienda a los estudiantes hacer esta matriz de congruencia en una figura como la 11.1 o una matriz.

Introducción	Marco teórico	Método	Resultados	Discusión

6. Referencias, bibliografía

Son las fuentes primarias utilizadas por el investigador para elaborar el marco teórico u otros propósitos; se incluyen al final del reporte, ordenadas alfabéticamente y siguiendo un estilo de publicaciones (APA, Harvard, Vancouver, etcétera).

7. Apéndices

Resultan útiles para describir con mayor profundidad ciertos materiales, sin distraer la lectura del texto principal del reporte o evitar que rompan con el formato de éste. Algunos ejemplos de apéndices serían el cuestionario utilizado, un nuevo programa computacional, análisis estadísticos adicionales, la descripción de un equipo utilizado, el desarrollo de una fórmula complicada, fotografías, etc. Se titulan y numeran en orden progresivo.

Reconocimientos y declaración de no conflicto de intereses

En los artículos suelen agregarse al final los agradecimientos y una declaración de que el autor o autores no tienen conflicto de intereses respecto a la autoría y la publicación del artículo. También se menciona, si aplica, la fuente de financiamiento del estudio.

¿Qué elementos contiene un reporte de investigación o informe de resultados en un contexto no académico?

Un reporte no académico contiene la mayoría de los elementos de un reporte académico:

1. Portada
2. Índice
3. Resumen ejecutivo (resultados más relevantes y casi todos presentados de manera gráfica)
4. Método
5. Resultados
6. Conclusiones[10]
7. Apéndices

Pero cada elemento se trata con mayor brevedad y se eliminan las explicaciones técnicas que no puedan ser comprendidas por los usuarios. Frecuentemente, el marco teórico y la bibliografía suelen omitirse del reporte o se agregan como apéndices o antecedentes, y cuando se incluyen son breves. Desde luego, lo anterior de ninguna manera implica que no se haya desarrollado un marco teórico, sino que algunos usuarios prefieren no confrontarse con éste en el informe de investigación. Hay usuarios no académicos que sí se interesan por el marco teórico y las citas bibliográficas y referencias. Para ilustrar la diferencia entre redactar un reporte académico y uno no académico, se presenta un ejemplo de introducción de un reporte no académico que, como se ve en el siguiente recuadro de ejemplo, es bastante sencillo, breve y no utiliza términos complejos.

[10] En los ambientes no académicos se usa el término "conclusiones" en lugar de discusión.

Ejemplo

Muestra de introducción de un reporte no académico

Calidad total

La Fundación Mexicana para la Calidad Total, A.C. (Fundameca), realizó una investigación por encuestas para conocer las prácticas, técnicas, estructuras, procesos y temáticas existentes en calidad total en nuestro país. La investigación, de carácter exploratorio, constituye el primer esfuerzo para obtener una radiografía del estado de los procesos de calidad en México. No es un estudio exhaustivo, sólo implica un primer acercamiento, que en los años venideros irá extendiendo y profundizando la fundación.

El reporte de investigación que a continuación se presenta tiene como uno de sus objetivos esenciales propiciar el análisis, la discusión y la reflexión profunda respecto de los proyectos para incrementar la calidad de los productos o servicios que ofrece México a los mercados nacional e internacional. Como nación, sector y empresa: ¿vamos por el camino correcto hacia el logro de la calidad total? ¿Qué estamos haciendo adecuadamente? ¿Qué nos falta? ¿Cuáles son los obstáculos a los que nos estamos enfrentando? ¿Cuáles son los retos que habremos de rebasar en la primera década del siglo? Éstas son algunas de las preguntas que estamos valorando y necesitamos responder. La investigación pretende aportar algunas pautas para que comencemos a contestar en forma satisfactoria dichos cuestionamientos.

La muestra de la investigación fue seleccionada al azar sobre la base de tres listados: 500 empresas de *Expansión*,[11] listado de la gaceta *Cambio Organizacional* y registro de empresas asistentes a las reuniones para constituir Fundameca. Se acudió a 184 empresas, de las cuales 60 no proporcionaron información. Dos encuestas fueron eliminadas por detectarse inconsistencias lógicas. En total se incluyeron 122 casos válidos.

Esperamos que sus comentarios y sugerencias amplíen y enriquezcan este proceso investigativo.

Fundameca
Dirección de Investigación

Recomendaciones para redactar un reporte de investigación

Los reportes de investigación requieren de escribir, revisar lo escrito una y otra vez, y editar, basándonos en nuestra propia crítica y los comentarios de los demás, hasta que estemos satisfechos con la claridad de nuestras explicaciones y narrativa (y que colegas o pares nos aprueben el documento). Desde luego, no hay escrito ni informe perfecto, pero aquellos documentos que se revisan cuidadosamente varias veces por el propio autor y otros investigadores, tienen una mayor probabilidad de resultar comprensibles, precisos, concisos y apegados a las reglas de publicación de los estilos aceptados por las comunidades académicas y profesionales, así como a las reglas del idioma en el cual fueron elaborados.

Algunas recomendaciones para redactar un informe de investigación son las siguientes:

- Trabaje siguiendo un boceto o bosquejo (inicie con un primer índice o índice tentativo).
- Al redactar un apartado o capítulo tenga en mente los demás (por ejemplo, si está redactando el método, concéntrese en éste pero considere el índice tentativo de los capítulos de resultados y discusión y lo que escribió en el marco teórico).
- Utilice la voz activa.
- Al redactar mezcle adecuadamente los tiempos (tal autor señala…, el instrumento utilizado fue…, esta investigación difiere de otras en…, la hipótesis nula se acepta, los resultados son consistentes con, la muestra fue…, se midió…).
- Elimine términos ambiguos, redundantes o innecesarios.
- Busque sinónimos para no duplicar palabras en una misma oración o párrafo.
- Comience cada capítulo de la tesis o parte del documento con un párrafo introductorio y señale la estructura, es decir, cómo está compuesto (aplica a marco teórico, método, resultados y discusión).
- Al final de cada parte, escriba oraciones o párrafos de transición para señalar lo que sigue en la próxima. Vincule entre apartados de una misma sección o capítulo y evite "saltos conceptuales".

[11] Similar al de las 500 de Fortune, pero exclusivamente en México.

- Vaya ligando párrafos paulatinamente.
- En los primeros párrafos que redacte de cada apartado, suprima la autocrítica, deje fluir su escritura. Ya que perciba que tiene fluidez, entonces revise críticamente.
- Utilice títulos o encabezados como guías para el lector.
- Asegúrese de que las secciones de un apartado están relacionadas entre sí.
- Al redactar el marco teórico o revisión de la literatura verifique que cada apartado o sección esté vinculada con el planteamiento (preguntarse: ¿este apartado está vinculado al planteamiento? Si no es así, es posible que ese apartado no venga al caso y tenga que eliminarse).
- Revise el índice y el manuscrito con otros colegas que sepan del tema, su profesor y compañeros. Incluso, es conveniente que una persona no especialista en el tema (pero con educación profesional) lea el documento para verificar la claridad.
- Cabe destacar que en reportes para publicarse, como los artículos de una revista científica, se desarrollan todos los elementos de manera muy concisa o resumida.
- En todo momento deben buscarse explicaciones directas, así como eliminar repeticiones, argumentos innecesarios y redundancia no justificada. En el lenguaje debemos ser muy cuidadosos y sensibles, no debemos utilizar términos despectivos refiriéndonos a personas con capacidades distintas, grupos étnicos diferentes al nuestro, etc.; para ello, es necesario consultar algún manual de los que se recomiendan más adelante.

En el centro de recursos en línea, el lector podrá encontrar en "Manuales auxiliares", un documento titulado "Indicaciones para la elaboración de un manuscrito original", elaborado por los departamentos de investigación de la Universidad Católica de Costa Rica y la Universidad de Celaya, que contiene diversas recomendaciones y lineamientos para escribir informes de investigación en distintos formatos; así como el capítulo 11 adicional: "Consejos prácticos para realizar investigación" (en Material complementario → Capítulos).

Centro de recursos en línea

¿Hay guías o manuales para elaborar un reporte de investigación?

Para desarrollar el reporte siempre debemos seguir un estilo de publicación, el cual nos indica cómo citar las referencias en el texto del informe e incluirlas en la bibliografía (lista de referencias), así como la manera de presentar tablas, figuras, ecuaciones y otros elementos. Es una forma que las comunidades académicas han desarrollado para construir estándares homogéneos en la redacción de reportes y dar crédito a las fuentes originales consultadas, así como comunicarle al lector dónde puede localizarlas.

Los principales estilos de publicación son:

1. APA (American Psychological Association)
2. Vancouver
3. Harvard
4. Chicago
5. MLA (Modern Language Association)

Cada uno tiene sus propias reglas y pueden encontrarse manuales sobre el estilo específico. Al inicio de su desarrollo algunos estilos se asociaron con ciertas disciplinas, pero hoy en día esto ha cambiado. Por ejemplo, el estilo APA es utilizado no solamente por asociaciones de psicólogos, sino también por administradores y agrupaciones de ingenieros químicos. Entonces, ¿cuál utilizar como guía al elaborar un informe o un artículo para una revista científica? En definitiva el que se utilice como norma en nuestra institución o escuela. En el caso de revistas y congresos, los editores o revisores indican el estilo (por ejemplo, en las instrucciones para los autores o equivalente). Cabe destacar que una vez elegido cierto estilo es requisito apegarse a éste; no pueden mezclarse estilos en un mismo reporte.

Centro de recursos en línea

Por cuestiones de espacio y sobre todo para mantener al máximo posible su actualización no serán tratados en la parte impresa de la obra, pero en el capítulo 11, "Consejos prácticos para realizar investigación", del centro de recursos que puede encontrar en: Material complementario → Capítulos,

www.elosopanda.com | jamespoetrodriguez.com

y descargarse, se presenta una tabla comparativa de los estilos Vancouver, APA, Harvard y Chicago en cuanto a cómo incluir las principales fuentes en la bibliografía (libros, capítulos, artículos, etc.). Por otro lado, se incluyen ligas a sitios en internet sobre los estilos (los mencionados y otros).

Asimismo, en dicho centro (en la sección "Manuales auxiliares"), el lector podrá encontrar y bajar a su dispositivo un manual sencillo titulado: "Introducción al estilo APA para citas y referencias" (que en la quinta edición se incluyó en el CD) y el ya conocido programa SISI (Sistema de Información para el Soporte a la Investigación) para capturar documentos y generar, incluir y organizar citas o referencias en el texto y al final en la bibliografía o listado de referencias, sobre la base del estilo APA (el software las coloca de manera correcta).

Adicionalmente, podrá localizar en el centro de recursos otro manual sencillo nombrado: "Introducción al estilo Vancouver para citas y referencias".

¿Qué recursos están disponibles para presentar el reporte de investigación?

Son hoy tantos los programas de dibujo, de gráficas, presentaciones y elaboración de documentos, que es imposible en este espacio comentarlos o siquiera nombrarlos. Use todos los que conozca y tenga acceso a ellos, recuerde que una presentación debe tener riqueza visual. En los documentos hay ciertas reglas que no podemos hacer a un lado, pero en la presentación el límite es nuestra propia imaginación.

¿Qué criterios o parámetros podemos definir para evaluar una investigación o un reporte?

Un buen reporte debe demostrar que la investigación respondió con claridad al planteamiento del problema, ese es el mejor parámetro. Hay reportes que han sido tan trascendentes que sus hallazgos y recomendaciones han revolucionado un campo y permanecido en el tiempo. Por ejemplo, Hernández Galicia (1970) realizó un estudio respecto a los efectos de los gases emitidos por los motores de combustión interna sobre la contaminación de la atmósfera de la Ciudad de México. Sus conclusiones y pronósticos llegan hasta nuestros días y desafortunadamente no se acataron sus recomendaciones. Y no es cuestión de clarividencia, sino de que un estudio científico hecho rigurosamente puede predecir realidades.

Una propuesta de parámetros o criterios para evaluar la calidad de un estudio cuantitativo y, consecuentemente, su informe, se presenta en el centro de recursos, en: Material complementario → Capítulos → Capítulo 10 "Parámetros, criterios, indicadores o cuestionamientos para evaluar la calidad de una investigación".

¿Con qué se compara el reporte de la investigación? ¿Y la propuesta o protocolo de investigación?

El reporte se contrasta con la propuesta o protocolo de la investigación, la que hicimos al inicio del proceso, que no se ha comentado en el libro, porque primero resultaba necesario conocer el proceso de investigación cuantitativa.

El protocolo se revisará en el centro de recursos, en: Material complementario → Capítulos → Capítulo 9, "Elaboración de propuestas cuantitativas, cualitativas y mixtas".

Resumen

- Antes de elaborar el reporte de investigación debe definirse a los usuarios, ya que el informe habrá de adaptarse a ellos.
- Los reportes de investigación pueden presentarse en un contexto académico o en un contexto no académico.
- Los usuarios y el contexto determinan el formato, la naturaleza y la extensión del informe de investigación.
- Las secciones más comunes de un reporte de investigación presentado en un contexto académico son: portada, índice, resumen, palabras clave, cuerpo del documento (introducción, marco teórico, método, resultados), discusión, referencias o bibliografía y apéndices.
- Los elementos más comunes en un contexto no académico son: portada, índice, resumen ejecutivo, método, resultados, conclusiones y apéndices.
- Centro de recursos en línea Los reportes deben seguir un estilo de publicaciones, los principales son: APA, Vancouver, Harvard y Chicago. En el centro de recursos puede descargar información relacionada en el capítulo 11 y los manuales sencillos de los estilos APA y Vancouver.
- Para presentar el reporte de investigación se pueden utilizar diversos apoyos o recursos.

Conceptos básicos

- Apéndices
- Contexto académico
- Contexto no académico
- Cuerpo del documento
- Discusión
- Estilo de publicación
- Introducción
- Marco teórico (Revisión de la literatura)
- Método
- Referencias
- Reporte de investigación
- Resultados
- Resumen
- Usuarios/receptores

Ejercicios

1. Elabore el índice de una tesis.
2. Centro de recursos en línea Localice un artículo de una revista científica mencionada en el apéndice 1 del centro de recursos y analice las secciones del artículo.
3. Desarrolle el índice del reporte de la investigación que ha concebido a lo largo de los ejercicios del libro.
4. Elabore una presentación de su tesis o de cualquier investigación realizada por usted u otra persona en un programa para tal efecto disponible en su institución (por ejemplo: Power Point, Prezi o Flash).

5.5.2. Elaboración de conclusiones fundamentadas y su relación con las hipótesis

Para el desarrollo del subtema **Elaboración de conclusiones fundamentadas y su relación con las hipótesis**, se consideran aquellos contenidos del libro *Metodología de la investigación* de **César Augusto Bernal Torres** que permiten comprender la formulación de las conclusiones a partir del análisis y discusión de los resultados, así como su articulación con los objetivos y las hipótesis de la investigación. Se priorizan los apartados relacionados directamente con las **conclusiones, el análisis de resultados y discusión, las hipótesis y los objetivos de la investigación**, debido a que estos contenidos permiten establecer la relación lógica entre los elementos del proceso investigativo.

N.°	Libros	Capítulo	Página	Contenido
1	*Metodología de la investigación* — César Augusto Bernal Torres	**Capítulo 7. Proceso de investigación científica — 7.16. Conclusiones**	**p. 228**	**Conclusiones.** Contenido principal para el desarrollo del subtema, por abordar directamente la elaboración de conclusiones dentro del proceso de investigación.
2	*Metodología de la investigación* — César Augusto Bernal Torres	**Capítulo 7. Proceso de investigación científica — 7.13. Análisis de resultados. Discusión**	**p. 220**	**Análisis de resultados y discusión.** Constituye el fundamento previo para elaborar conclusiones a partir de los resultados obtenidos en la investigación.
3	*Metodología de la investigación* — César Augusto Bernal Torres	**Capítulo 7. Proceso de investigación científica — 7.7. Hipótesis de la investigación**	**p. 136**	**Hipótesis de la investigación.** Aporta el contenido necesario para establecer la relación entre las hipótesis formuladas y las conclusiones derivadas de los resultados.
4	*Metodología de la investigación* — César Augusto Bernal Torres	**Capítulo 7. Proceso de investigación científica — 7.3. Objetivos de la investigación**	**p. 97**	**Objetivos de la investigación.** Permite relacionar las conclusiones con los objetivos planteados al inicio del proceso investigativo.
5	*Metodología de la investigación* — César Augusto Bernal Torres	**Capítulo 7. Proceso de investigación científica — 7.11. Procesamiento de la información. Datos**	**p. 198**	**Procesamiento de la información y datos.** Contenido complementario que permite comprender la base de información sobre la cual posteriormente se realizan el análisis, la discusión y la elaboración de conclusiones.
6	*Metodología de la investigación* — César Augusto Bernal Torres	**Capítulo 7. Proceso de investigación científica — 7.12. Modelos de procesamiento de datos con el uso de herramientas estadísticas**	**p. 203**	**Procesamiento de datos mediante herramientas estadísticas.** Contenido complementario para investigaciones en las que las conclusiones se fundamentan en resultados obtenidos mediante procedimientos estadísticos.

5.5.2. Elaboración de conclusiones fundamentadas y su relación con las hipótesis

Para el desarrollo del subtema **Elaboración de conclusiones fundamentadas y su relación con las hipótesis**, se consideran aquellos contenidos de la obra *Metodología de la investigación* de **Roberto Hernández Sampieri** que permiten comprender la formulación y comprobación de las hipótesis, el análisis de los datos, la interpretación de los resultados y la preparación del reporte de investigación. Se priorizan los contenidos que permiten establecer la relación entre **hipótesis, análisis de datos, resultados y conclusiones**, considerando especialmente los apartados dedicados a la formulación de hipótesis, la prueba de hipótesis, el análisis de datos cuantitativos y la presentación de los resultados.

N.°	Libros	Capítulo	Página	Contenido
1	*Metodología de la investigación* — Roberto Hernández Sampieri	**Capítulo 6. Formulación de hipótesis**	**p. 102**	**Hipótesis de la investigación.** Concepto, tipos de hipótesis, utilidad y función dentro del proceso de investigación.
2	*Metodología de la investigación* — Roberto Hernández Sampieri	**Capítulo 6. Formulación de hipótesis**	**p. 102**	**¿Qué ocurre cuando no se aporta evidencia a favor de las hipótesis de investigación?** Contenido directamente relacionado con la interpretación de los resultados y su relación con las hipótesis.
3	*Metodología de la investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**p. 270**	**Análisis de datos cuantitativos.** Procesamiento y análisis de los datos obtenidos para generar resultados que permitan fundamentar las conclusiones.
4	*Metodología de la investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**p. 299**	**Paso 5: analizar mediante pruebas estadísticas las hipótesis planteadas.** Procedimiento relacionado directamente con el análisis y contraste de las hipótesis a partir de los datos obtenidos.
5	*Metodología de la investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**p. 299**	**¿En qué consiste la prueba de hipótesis?** Fundamento para establecer la relación entre la evidencia obtenida mediante los datos y las hipótesis formuladas.
6	*Metodología de la investigación* — Roberto Hernández Sampieri	**Capítulo 10. Análisis de datos cuantitativos**	**p. 327**	**Paso 7: preparar los resultados para presentarlos.** Organización y preparación de los resultados obtenidos en el proceso de análisis.

7	*Metodología de la investigación* — Roberto Hernández Sampieri	**Capítulo 11. El reporte de resultados del proceso cuantitativo**	**p. 336**	**El reporte de resultados del proceso cuantitativo.** Presentación y comunicación de los resultados de la investigación como parte del cierre del proceso investigativo.
8	*Metodología de la investigación* — Roberto Hernández Sampieri	**Capítulo 12. El inicio del proceso cualitativo: planteamiento del problema, revisión de la literatura, surgimiento de las hipótesis e inmersión en el campo**	**p. 356**	**¿Qué papel desempeñan las hipótesis en el proceso de investigación cualitativa?** Contenido complementario para comprender el papel de las hipótesis en investigaciones cualitativas.

CAPÍTULO

7

Proceso de investigación científica

www.FreeLibros.me

Objetivos de aprendizaje

Al terminar de estudiar este capítulo, usted será capaz de:

1. Explicar la diferencia entre la visión lineal o causal y la visión sistémica o circular del proceso de investigación científica, en el método general o tradicional.
2. Identificar temas de investigación en el campo de su disciplina o campo del conocimiento.
3. Plantear el problema de investigación sobre un tema de interés.
4. Formular los objetivos (general y específicos) de la investigación.
5. Justificar y delimitar la investigación.
6. Elaborar el respectivo marco teórico de la investigación.
7. Definir el tipo de estudio por realizar.
8. Plantear la hipótesis de investigación cuando sea necesario.
9. Definir la población objeto del estudio y la(s) técnica(s) por utilizar en la obtención de la información.
10. Estimar el tamaño de muestra requerida para validar la información que se obtenga del estudio.
11. Definir el procedimiento para llevar a cabo la investigación.
12. Describir la forma como se procesará la información y se presentarán los resultados de la respectiva investigación.
13. Elaborar el cronograma de actividades y el presupuesto para el desarrollo de la investigación.
14. Presentar la bibliografía consultada en la elaboración de la propuesta o anteproyecto de investigación.

www.FreeLibros.me

No obstante la diversidad y cantidad de métodos de investigación, el objetivo principal de este capítulo es ejemplificar el desarrollo del proceso de investigación científica a través del *método general*.

Es usual, como reconocen estudiosos y profesores de metodología de la investigación científica, que quienes se inician en el campo de la investigación enfrentan, muy confusos, el uso del método de investigación y tienden a seguir, de manera rígida y mecánica, tal proceso como una receta igualmente aplicable a cualquier tema. Haciendo a un lado que aunque el método puede ser el mismo, cada proyecto tiene sus particularidades, por lo cual es necesario y pertinente adecuar el método al respectivo proyecto.

De acuerdo con Wallace (1976), el desarrollo de la investigación científica debe ser un proceso *circular*, *no lineal*, de interdependencia entre los elementos o aspectos constitutivos del método científico general que intervienen en la dinámica de la generación del conocimiento válido.

La Figura 7.1 muestra la dinámica del proceso de investigación científica propuesta por Wallace, que sirve como fundamento para el desarrollo del contenido de este capítulo.

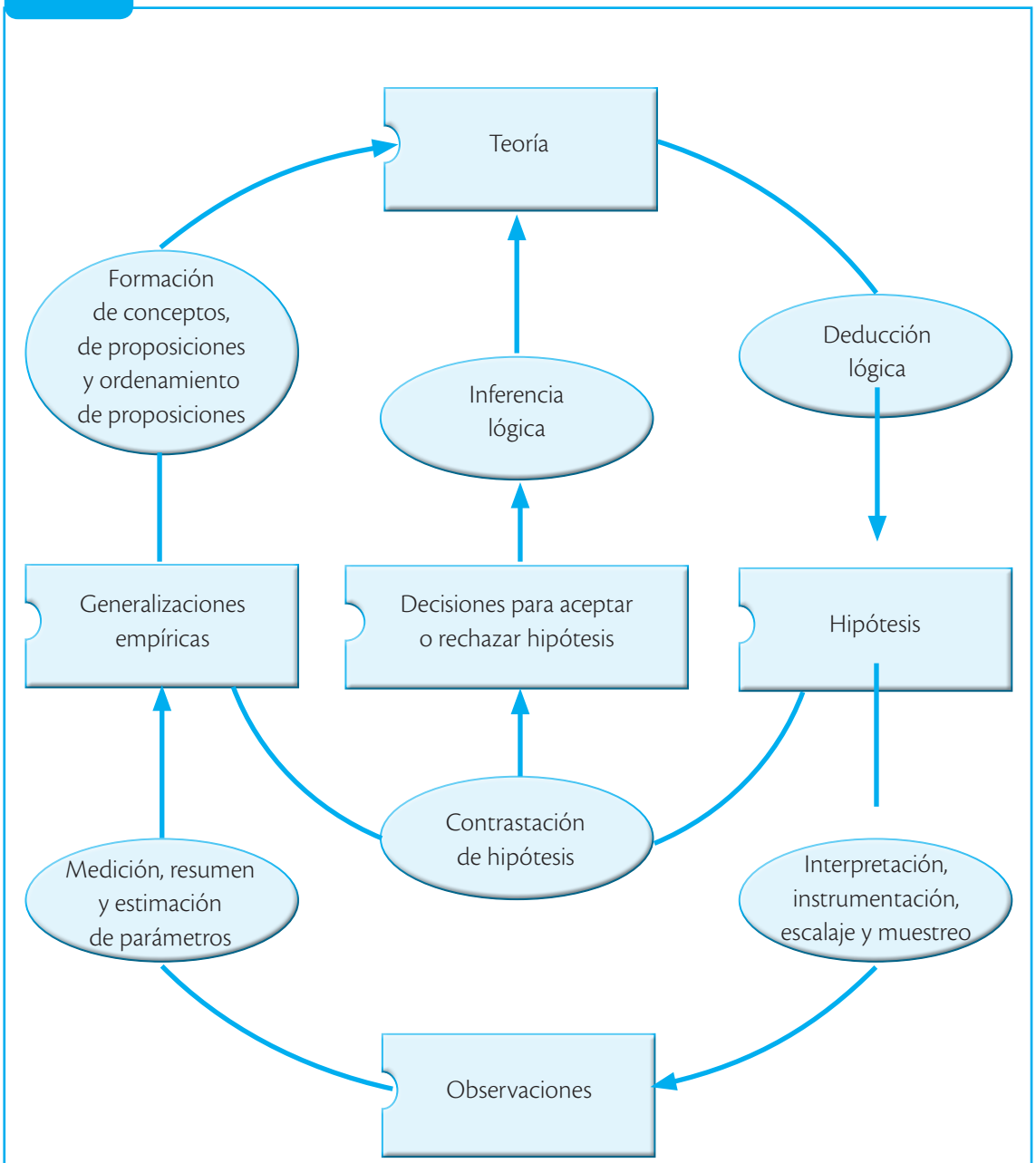
Respecto a los componentes o elementos que conforman el proceso de investigación, en su libro *Los elementos de la investigación*, Hugo Cerda (1998) considera que, al analizar cuáles son los componentes básicos de un proceso de investigación, existen numerosas alternativas, propuestas y sugerencias de los investigadores. Sin embargo, según este autor, en todas las propuestas hay puntos de coincidencia, entre los que se mencionan como los componentes básicos del proceso de investigación científica los siguientes:

1. La selección del tema de investigación.
2. El problema de investigación.
3. Objetivos generales y específicos.
4. El marco teórico.
5. Los recursos humanos, institucionales, técnicos y económicos.
6. La caracterización y delimitación de la población.
7. La selección de los métodos, las técnicas y los instrumentos de investigación.
8. La fuente de datos.
9. El trabajo de campo y trabajo de gabinete.
10. La tabulación, el análisis y la interpretación de datos.

La Figura7.2 ilustra los componentes del proceso de investigación, resultado de la discusión presentada en los capítulos anteriores, el planteamiento de Cerda y la experiencia del autor en los cursos de metodología de la investigación. Así, el proceso de investigación científica para efectos de este libro es un proceso circular de interacción entre sus diferentes componentes, es decir, es un sistema en forma de espiral, en el que, a medida que va desarrollándose cada componente se soporta en los previamente definidos, manteniendo el criterio de unidad y coherencia entre ellos.

www.FreeLibros.me

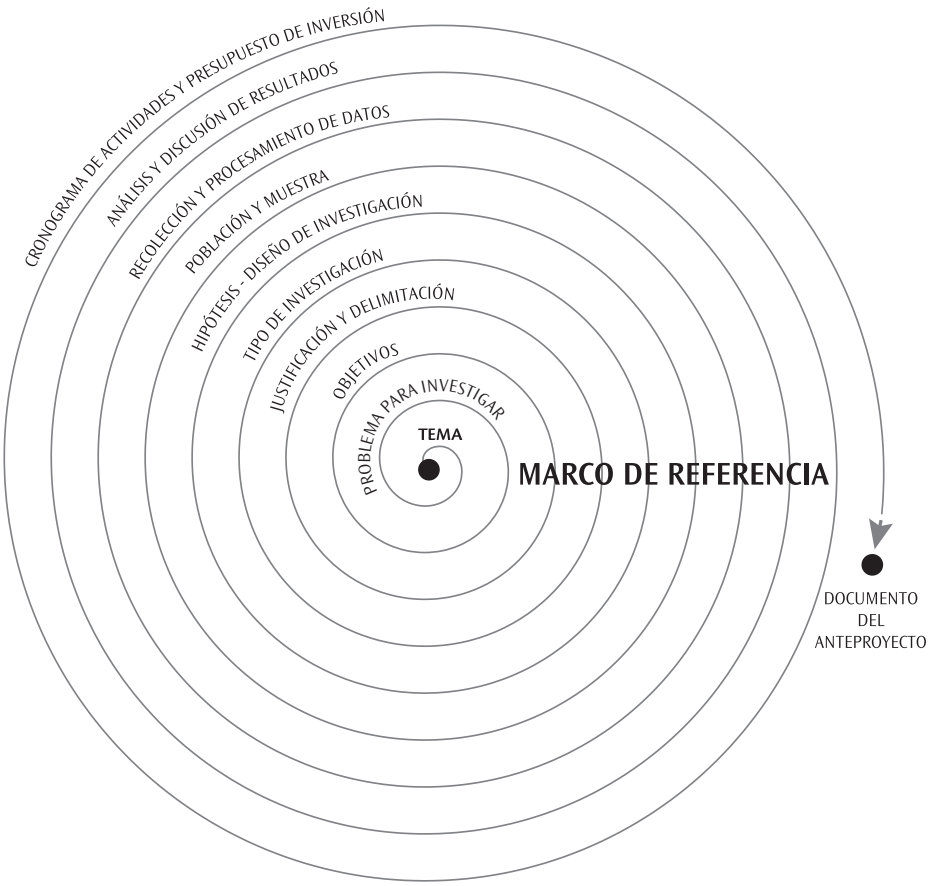
FIGURA 7.1 **Dinámica del proceso de investigación científica**



Nota: los componentes de información van en rectángulos; los controles metodológicos, en óvalos, y las transformaciones de información se indican con flechas (Wallace, 1976).

www.FreeLibros.me

FIGURA 7.2 **Componentes del proceso de investigación científica**

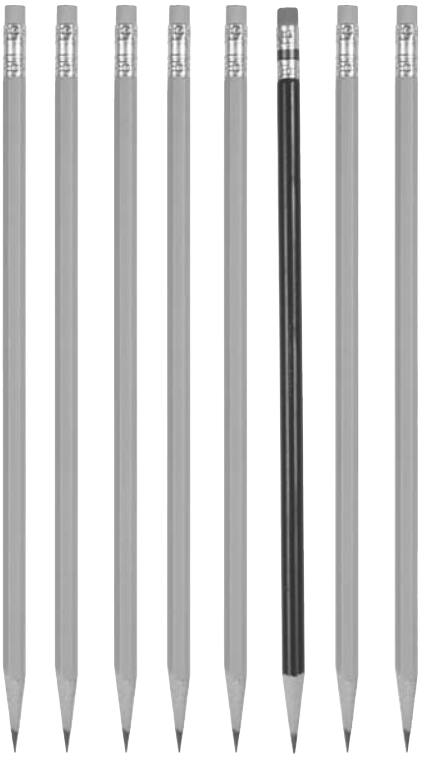


En síntesis, el proceso de investigación es un sistema constituido por varios componentes que a medida que va desarrollándose, cada componente recibe influencia del anterior, pero, a la vez, es seguido e influido por otro. En cada fase o etapa se desarrolla un componente, aunque no todos ellos siguen una secuencia de etapas.

Por ejemplo, el marco de referencia se desarrolla desde el inicio hasta el final de la investigación, mientras que otros componentes (los objetivos y las hipótesis) se definen en un determinado momento, a pesar de que pueden modificarse durante la investigación. En el caso de la población objeto de estudio, ésta debe definirse antes de iniciar la fase de recolección de la información.

A continuación se describen los componentes sugeridos en este libro para el desarrollo de un proceso de investigación científica. Cada uno comienza con una fundamentación teórica básica; además, se desarrollan algunos ejemplos que ilustran el tema y se finaliza con algunas preguntas de repaso y análisis.

www.FreeLibros.me



7.1 Tema de investigación

En el *método general* de la investigación científica, es usual que la investigación comience por despertar interés por un tema en especial. Por tal motivo, a continuación se muestran varios aspectos para tener en cuenta al momento de elegir o definir un tema de investigación. Estos son:

- Búsqueda y definición del tema.
- Criterios para considerar la pertinencia del tema.
- Medios para categorizar la relevancia del tema.
- Título del tema por investigarse.

El proceso de investigación científica usualmente comienza con el interés por un tema de investigación. Pero ¿de dónde surge el tema, qué características debe cumplir para ser considerado un tema de investigación y quiénes lo consideran como tal? Éstos interrogantes se responden a continuación.

7.1.1 Búsqueda y definición del tema

El *tema* es la idea general del campo del conocimiento de una disciplina, en el cual hay interés para realizar una investigación. En el caso de los trabajos de grado, es necesario que el tema de investigación sea pertinente con los contenidos de la disciplina o profesión a la que se optará al respectivo grado, es decir, si el trabajo de grado es para optar al título de administración, los temas de investigación han de referirse al campo de la administración; si es para optar al título de economía, los temas han de ser del campo de la economía, y si es para optar al título de psicología, los temas han de ser del campo de la psicología, etcétera.

Por ejemplo, en el caso de la administración, los temas de investigación deben contemplar cualquiera de los siguientes campos: la teoría de la administración, el proceso administrativo (planeación, organización, dirección y control) o las áreas funcionales (producción, mercadeo, desarrollo humano, finanzas, gerencia, etcétera).

Para el caso de la psicología algunos campos del conocimiento son: psicología clínica, psicología educativa, psicología evolutiva, psicología de la adolescencia, psicología organizacional, etcétera. De estos campos emergen subcampos, que pueden ser cada vez más específicos; lo mismo ocurre con los temas de investigación. Inicialmente surgen ideas generales como investigar en "psicología organizacional"; luego uno se pregunta en forma específica qué quiere investigar en este subcampo. Una respuesta sería la "motivación en el trabajo", por lo que la pregunta siguiente sería: ¿motivación en el trabajo a nivel general, en un contexto en particular o en una empresa específica? Así se delimita el tema.

www.FreeLibros.me

Los temas de investigación surgen de diversas formas, y para descubrirlos se necesita interés por la investigación y una actitud dinámica y reflexiva respecto a los diferentes conocimientos inherentes a cada profesión.

Entre las distintas formas generadoras de temas de investigación se mencionan:

Lectura reflexiva y crítica Se deben consultar fuentes como libros, revistas especializadas y demás documentos que plantean reflexiones sobre la respectiva disciplina o que, siendo de otra, aportan algo a la disciplina de interés. Por ejemplo, las revistas especializadas en psicología podrían plantear reflexiones sobre un tema específico que se transfiera a disciplinas como la administración y la economía.

Participación activa Al asistir y participar en conferencias, congresos, discusiones, eventos académicos y demás formas de exposición y reflexión de temas, los expositores y otras personas del público plantean reflexiones que llegan a convertirse en temas objeto de investigación.

Experiencia individual Cada persona, sin proponérselo, tropieza muchas veces con interrogantes acerca de su disciplina en particular o sobre el quehacer de la ciencia en general, que bien podrían llegar a ser temas de investigación.

Práctica profesional Este ejercicio genera temas de investigación, cuando se es un profesional crítico y con deseos de aportar a la disciplina propia.

Aula de clase Independientemente de los recursos didácticos que se utilicen en el aula, si el alumno es activo y reflexivo, se encontrará con muchísimos interrogantes que, bien formulados, serían temas de investigación.

Centros de investigación Muchas universidades tienen centros de investigación donde se generan y desarrollan líneas de investigación en las respectivas disciplinas.

Organismos interesados en la investigación Muchas instituciones desean desarrollar proyectos de investigación tendientes a resolver problemas de su competencia. Es muy importante estar atentos a estas instituciones porque, a la vez que tienen definidos los temas de interés para la investigación, aportan recursos financieros, humanos y técnicos y son un excelente medio para aprender a investigar con el apoyo de expertos.

Los profesores Algunos trabajadores docentes están interesados en investigar un tema determinado y requieren apoyo de estudiantes para desarrollarlo.

Según Muñoz Giraldo et al. (2001), son fuentes de tema y problema de investigación:

- La experiencia.
- Los vacíos del conocimiento en el campo de la disciplina.
- Los resultados contradictorios de otras investigaciones.
- La necesidad de explicaciones acerca de los hechos o los fenómenos.
- La incoherencia entre la teoría y la práctica en un tema determinado.

www.FreeLibros.me

- La necesidad de verificar, descubrir, crear y solucionar dificultades.
- La diversidad de teorías sobre un tema o campo del conocimiento.
- El conocimiento sobre un tema a partir de resultados de investigaciones que pueden replicarse o generar nuevas preguntas.
- Los aportes y discusiones de otros investigadores con las mismas inquietudes.
- Las orientaciones filosóficas que modelan los intereses, así como las formas de pensar y de producir conocimiento.
- Ideologías culturales, valores, conflictos sociales, tecnológicos y morales, típicos de un contexto específico.
- Cuestiones o inquietudes indicadas por comunidades científicas o por la sociedad en general.
- La reflexión sobre la propia práctica, las reacciones de los colegas y la crítica argumentada.

Como complemento de lo anterior, y acorde con Hugo Cerda, algunas ideas que pueden hacer surgir temas de investigación serían (Cerda, 1988):

- Una necesidad que debe satisfacerse.
- Una causa que hay que determinar, descubrir, precisar o explicar.
- La necesidad de conocer la relación entre fenómenos, objetos o situaciones.
- Una dificultad que debe ser superada, identificada o explicada.
- La necesidad de determinar la existencia, vigencia y viabilidad de un objeto o un proyecto.
- La identificación de un fenómeno o un aspecto que se considera importante o vigente en un momento determinado.
- La comprensión de una cosa, de sus relaciones, de sus efectos o de sus valores.
- La clasificación o tipología necesarias para plantear o comprender un fenómeno o una situación.
- La determinación de la propiedad de un fenómeno, de una actividad o de un conjunto de personas, con el propósito de definirlos, describirlos, analizarlos, etcétera.
- La descripción de un objeto o de un fenómeno, con el propósito de identificar, definir o analizar las características o propiedades de ese objeto o fenómeno.
- El análisis de un hecho particular, por su novedad y situación particular.

Aunque la lista es un buen ejemplo de la abundancia de temas de investigación, hay que recordar que una época como la que vivimos, caracterizada por la incertidumbre, el cambio, los desequilibrios, el estrés, la competitividad y la globalización, y en donde el conocimiento se vuelve rápidamente obsoleto, se ha convertido en un escenario propicio de temas de investigación; por otro lado, y puesto que nuestros países tienen poca tradición en investigación, debe tenerse en cuenta que son innumerables los problemas que esperan ser investigados.

www.FreeLibros.me

No obstante, se escuchan con frecuencia entre los estudiantes de últimos semestres expresiones como: "No tengo tema de investigación para mi trabajo de grado", "¿Sobre qué tema puedo investigar?", etcétera.

En general, existe una gran diversidad de medios para generar ideas y desarrollar propuestas de investigación; una vez conocidos éstos, resulta sencillo obtener temas para la realización de un trabajo de grado.

7.1.2 Criterios para considerar la pertinencia del tema

No existen criterios únicos para considerar un tema como fuente de investigación. Sin embargo, a continuación se plantean algunos que deben tenerse en cuenta:

- *Novedad*: en esencia porque el tema no se ha tratado o lo ha sido muy poco, o cuando se propone una nueva forma de abordar un problema o una situación.
- *Contraste*: es un tema que busca contrastar resultados de investigaciones anteriores en otros contextos.
- *Necesidad e importancia*: una situación determinada merece estudiarse por los argumentos que se exponen sobre la necesidad y la importancia de tratar el tema.
- *Resolución*: que el tema contribuya a resolver un problema específico.
- *Concreción y pertinencia*: esto es muy importante porque en muchas ocasiones se proponen temas demasiado generales que, además, no son pertinentes para determinado campo y disciplina del saber; por ejemplo, los componentes de la contaminación ambiental es un tema general y, para el caso de las ciencias económico-administrativas, es uno difícil de abordar, porque si lo que se desea estudiar son los componentes químicos, el estudiante no tendrá formación para ello. Otro ejemplo es conocer el proceso de aprendizaje de los directivos en el ejercicio de su papel gerencial; aunque específico, no es un tema para administradores ni economistas, y sí para psicólogos, pedagogos o psicopedagogos.
- *Lineamientos*: que responda a los lineamientos de investigación de la institución académica donde va a realizarse la investigación.

En fin, son variados los criterios para darle categoría de tema de investigación a una idea, pero éstos son los más usados.

7.1.3 Medios para categorizar la relevancia del tema

En el caso de proyectos de investigación para trabajos de grado, tres medios categorizan la relevancia de un tema de investigación:

Lectura sobre el tema al cual se refiere la idea Cuando surge una idea, es básico sondear qué se ha escrito sobre el tema y su importancia en el momento actual. Con esta información podría tenerse un concepto sobre la pertinencia o no de estudiar el tema.

www.FreeLibros.me

Expertos en el tema En los distintos campos del conocimiento, hay personas que saben del tema y pueden orientar acerca de su pertinencia y sugerir ajustes o cambios o, en otros casos, descartar el tema (porque ya ha sido trabajado, la idea es muy vaga, no pertenece al área de estudio, etcétera).

Coordinadores del área de trabajo de grado o centros de investigaciones Los coordinadores de investigación están actualizados en temas de estudio en su respectiva disciplina, cuentan con acceso a información relacionada en otras instituciones y tienen como misión guiar a los interesados en investigación y, por tanto, apoyarlos para definir sus temas de investigación.

7.1.4 Título del tema que se va a investigar

Definida la idea o el tema específico de interés para la investigación, es necesario condensarlo (sintetizarlo) en una frase que exprese la esencia de la idea o el tema que va a investigarse, la cual se denomina título del *estudio* o proyecto de investigación.

En el caso de los trabajos de grado (no sucede lo mismo con un libro), el título debe ser general, en cuanto recoge la esencia del tema que va a tratarse, pero específico en cuanto debe referirse al problema objeto de investigación.

El título debe demostrar el tema y, en particular, el problema que va a investigarse, que igualmente debe reflejarse en todo el proceso del desarrollo del estudio; por tanto, no se aconseja poner títulos generales, sino más bien específicos, como se muestra en los ejemplos 7.1 y 7.2 que se presentan a continuación. Además, el título puede modificarse durante el desarrollo de la investigación.

En síntesis, la Figura 7.3 ilustra los aspectos relevantes relacionados con el tema de investigación con centro en el interés por el tema que va a investigarse y del cual se desprenden sus principales fuentes, los criterios que deben utilizarse para su respectiva categorización y validación, así como su concreción en un título (a partir del tema se define el título de la investigación) para pasar al siguiente elemento del proceso de la investigación que es planteamiento del problema de investigación.

EJEMPLO 7.1 **Títulos de trabajos de investigación en el campo de los negocios y la economía**

- Si el interés es conocer el estado actual de la industria en un país, pero sólo está disponible información en empresas del sector financiero localizadas en una ciudad, el título podría ser: *Análisis de la situación actual de las empresas del sector financiero localizadas en Anacin.*
- Si el interés es estudiar la motivación en las empresas nacionales, el título de un trabajo de grado sería: *La motivación de personal en el sector hotelero de Calama.*

www.FreeLibros.me

- Si el interés es estudiar la contaminación ambiental causada por las empresas nacionales, el título de un trabajo de grado podría ser: *El manejo ambiental por parte de las empresas del sector de las artes gráficas*.
- Si el interés es estudiar estrategias de marketing, el título de un trabajo de grado sería: *Estrategias de marketing de las empresas nacionales frente a las utilizadas por las empresas extranjeras establecidas en el país*.
- Si el interés es estudiar el impacto de la tasa de cambio en las tasas de interés, el título podría ser: *Estudio evaluativo del impacto de la tasa de cambio en las tasas de interés durante los dos últimos años*.
- Si el interés es analizar el problema de desempleo en su país, un título de proyecto de investigación sería: *Análisis del desempleo en TAQUIA durante los últimos diez años*.
- Si el interés es estudiar el manejo de la política monetaria durante un determinado período en su país, el título de un trabajo de grado podría ser: *Análisis del impacto de la política monetaria del gobierno X en la actividad económica del país durante el período 2000-2005*.
- Si el interés es estudiar la inversión extranjera en su país, el título para un proyecto de investigación sería: *Incidencia de la inversión extranjera en el desarrollo industrial del país durante los últimos diez años*.
- Si el interés es estudiar el problema de la pobreza en su país o región, el título para una propuesta de investigación sería: *Análisis de las causas y consecuencias de la pobreza en x durante el período 2000-2010*.
- Si el interés es estudiar la ética de los profesionales de la contaduría pública, un título para un trabajo de grado sería: *Análisis de la actitud ética de los profesionales de la contaduría pública en el desarrollo de su profesión*.
- Si el interés es estudiar la función de la información contable en la administración de las organizaciones, el título para una posible investigación podría ser: *Importancia de la información contable en la administración efectiva de las organizaciones*.
- Si el interés es estudiar el costeo por actividades, el título de un proyecto de investigación sería: *Ventajas del sistema de costeo basado en actividades respecto del sistema de costeo tradicional*.
- En el campo de la ingeniería industrial, si el interés es estudiar la aplicación de la tecnología a los procesos productivos, un título para una propuesta de investigación podría ser: *Aplicaciones tecnológicas sistematizadas en el manejo de la producción en las empresas del sector manufacturero en la ciudad QLMT*.
- Si el interés es estudiar la producción limpia en las empresas, un título de un proyecto de investigación sería: *Diseño de un modelo de producción limpia para las empresas del sector de las artes gráficas*.
- Si el interés es estudiar el tema de la calidad, un posible título para investigación sería: *Análisis de los programas de calidad en el contexto de las empresas del sector manufacturero en QLMT*.

www.FreeLibros.me

FIGURA 7.3 **El proceso de investigación. Interés por el tema que va a investigarse**

www.FreeLibros.me

EJEMPLO 7.2 **Títulos de trabajos de investigación en educación y psicología**

- Cuando el interés es estudiar estrategias pedagógicas en la educación, el título del trabajo sería: *Evaluación del impacto de las estrategias pedagógicas utilizadas por los docentes de la educación media en el rendimiento académico de los estudiantes.*
- Si el interés es la calidad de la educación en la zona rural, un probable título sería: *Diagnóstico sobre calidad de la educación básica en la zona rural del departamento o estado JPTCP.*
- Un estudio sobre formación en valores llevaría el título: *Los valores éticos y sus implicaciones en la formación ética del estudiante.*
- Para un estudio sobre rendimiento académico, un título sería: *Análisis de la relación entre estrategias pedagógicas de enseñanza y estilos de aprendizaje en los estudiantes.*
- Un estudio sobre problemas intrafamiliares de las parejas se titularía: *Pautas de comportamiento en parejas con problemas intrafamiliares.*
- Un estudio sobre estrés en la familia llevaría el título: *Análisis de factores generadores de estrés y estrategias para el manejo del mismo en familias urbanas en San Salvador.*
- Un estudio sobre conductas suicidas de los adolescentes podría titularse: *Influencia del ambiente social, familiar y académico en la conducta suicida de los adolescentes en la ciudad AMTLK.*

RESUMEN

Con base en los planteamientos del capítulo 6, en la actualidad existe diversidad de métodos para realizar proyectos de investigación, razón por la cual los contenidos de este capítulo se orientaron a ejemplificar el desarrollo del proceso de investigación científica según el método general, que es uno de los diversos métodos de investigación, pero el más utilizado para tal efecto. Al respecto se precisó que quienes nos iniciamos en el campo de la investigación solemos tener confusión respecto al uso del método de investigación y tendemos a seguir, de manera rígida y mecánica, tal proceso como una receta igualmente aplicable a cualquier tema. Se nos dificulta entender que aunque el método puede ser el mismo, cada proyecto tiene sus particularidades, por lo cual es necesario y pertinente adecuar el método al respectivo tema del proyecto.

También se señaló la importancia de concebir el proceso de investigación científica como un proceso circular de interacción entre sus diferentes componentes, es decir, es un sistema en forma de espiral, en donde, a medida que va desarrollándose cada componente, el mismo se soporta en los componentes previamente definidos, manteniendo el criterio de unidad y coherencia entre todos ellos.

www.FreeLibros.me

En cuanto a los componentes básicos de un proceso de investigación, es importante recordar que existen numerosas alternativas, propuestas y sugerencias de los investigadores.

Sin embargo, hay coincidencia en que en el método general los componentes más usuales son los siguientes: seleccionar el tema, plantear el problema de la investigación, formular los objetivos generales y específicos, justificar y delimitar el tema, construir el marco teórico, definir el tipo de estudio por realizar, formular la hipótesis (cuando sea necesario), plantear el diseño experimental (cuando sea necesario), definir las estrategias metodológicas para el desarrollo de la investigación, realizar el trabajo de campo, procesar la información y analizar los resultados, redactar las conclusiones y el informe final.

En investigación, cuando se utiliza el *método general*, es usual comenzar por la definición del tema de interés para el investigador o equipo. Al respecto, y los aspectos para tomar en cuenta al momento de esta definición son: las fuentes de donde se obtendrá el tema, los criterios y los agentes que determinan sobre la pertinencia y relevancia del mismo. En el caso de los trabajos de grado, el tema de investigación ha de ser pertinente y coherente con los campos de estudio de la respectiva disciplina.

Una vez definida la idea o el tema específico de interés para la investigación, ésta se sintetiza en una frase que exprese la esencia de la misma, la cual da origen así al respectivo *título* del estudio o proyecto de investigación, título que puede modificarse durante el desarrollo de la investigación.

EJERCICIOS DE REPASO Y ANÁLISIS

1. ¿Por qué es importante tener ideas o temas de investigación?
2. Elabore una lista de temas que considere que ya hayan sido objeto de investigación en el campo de su disciplina o profesión.
3. Elabore una lista de temas sobre los cuales le interesaría profundizar o investigar en su campo o profesión.
4. Con base en la lista de temas del punto 3, referencie libros, capítulos, artículos, conferencias que haya leído sobre cada tema.
5. Con base en la lista de temas del punto 3, referencie las consultas que haya hecho con expertos en el respectivo tema y exponga una idea general para cada tema, consultando la opinión del experto.
6. Con base en la lista de temas del punto 3, explique la importancia de cada tema objeto de investigación.
7. En grupo, analicen la lista de temas de investigación del punto 3 en cuanto a claridad, pertinencia y especificidad.
8. En grupo, elaboren otra lista de temas de investigación que complemente la lista del punto 3.

www.FreeLibros.me

7.2 Problema de la investigación

Para que una idea sea objeto de investigación, debe convertirse en ***problema de investigación***. Ahora, en investigación, problema es todo aquello que se convierte en objeto de reflexión y sobre el cual se percibe la necesidad de conocer y, por tanto, de estudiar.

En este sentido, problema no es algo disfuncional, molesto o negativo, sino todo aquello que incite a ser conocido, pero teniendo en cuenta que su solución sea útil, es decir, buscando una respuesta que resuelva algo práctico o teórico. Por esto, a este modelo de investigación, además de ser conocido como modelo general, también suele denominársele *modelo pragmático*.

Una vez definidos el tema y el título de la *propuesta de investigación* se procede a plantear el problema de investigación, entendiendo como problema de investigación la situación, el fenómeno, el evento, el hecho u objeto del estudio que se va a realizar.

Arias Galicia (1991) considera que "en investigación no es suficiente visualizar un problema, es necesario plantearlo adecuadamente" (p. 50). Los especialistas en investigación enfatizan en la necesidad de un buen planteamiento del problema; para ellos, si esto se logra, la mitad del problema se ha solucionado. En este sentido, Briones (1985) afirma que "el planteamiento de un problema es la fase más importante de todo el proceso de investigación" (p. 39).

Plantear el problema de investigación significa enunciar y formular el problema, aspectos que se definen a continuación.

7.2.1 Enunciar el problema

Enunciar un problema de investigación consiste en presentar, mostrar y exponer las características o los rasgos del tema, situación o aspecto de interés que va a estudiarse, es decir, describir el estado actual del problema.

En general, enunciar un problema es contar lo que está pasando en relación con una situación, con una persona o con una institución; es narrar los hechos que caracterizan esa situación, mostrando sus implicaciones y soluciones.

Enunciar un problema requiere precisar la naturaleza y las dimensiones del mismo, en detalle y con precisión. Asimismo, se deben ambientar todas las características que enmarcan el problema; también hay que comenzar por narrar los antecedentes de la situación de estudio, así como incluir y mostrar los hechos, las relaciones y las explicaciones que sean importantes en la caracterización del problema. Igualmente, hay que contemplar tanto el problema como los elementos conectados con él (Tamayo, 2002).

www.FreeLibros.me

Tamayo (2002), citando a Van Dalen , sugiere tener en cuenta los siguientes aspectos al momento de plantear o definir un problema de investigación:

1. Reunir los hechos en relación con el problema (qué está pasando).
2. Determinar la importancia de los hechos.
3. Identificar las posibles relaciones entre los hechos que pueden indicar la causa de la dificultad.
4. Proponer explicaciones para conocer la causa de la dificultad y determinar su importancia en el problema.
5. Encontrar, entre las explicaciones, aquellas relaciones que permitan adquirir una visión amplia de la solución del problema.
6. Hallar relaciones entre los hechos y las explicaciones.
7. Analizar los supuestos en los que se apoyan los elementos identificados.

Para lograr definir o describir bien el problema hay que poseer un conocimiento previo sobre la situación que se va a estudiar, lo cual habrá de reflejarse en un conocimiento general que debe tener el investigador sobre el objeto o sujeto que se va a estudiar y sobre investigaciones específicas adelantadas en el asunto, así como algunas experiencias personales.

En síntesis, enunciar un problema es presentar una *descripción* general de la situación objeto de estudio.

7.2.2 Formular el problema

Un problema se *formula* cuando el investigador dictamina o hace una especie de *pronóstico* sobre la situación problema. En lugar de hacerlo con afirmaciones, este pronóstico se plantea mediante la formulación de preguntas orientadas a dar respuesta al problema de la investigación (Méndez, 1995).

Una adecuada formulación de un problema de investigación implica elaborar dos niveles de preguntas. La pregunta general debe recoger la esencia del problema y, por tanto, el título del estudio.

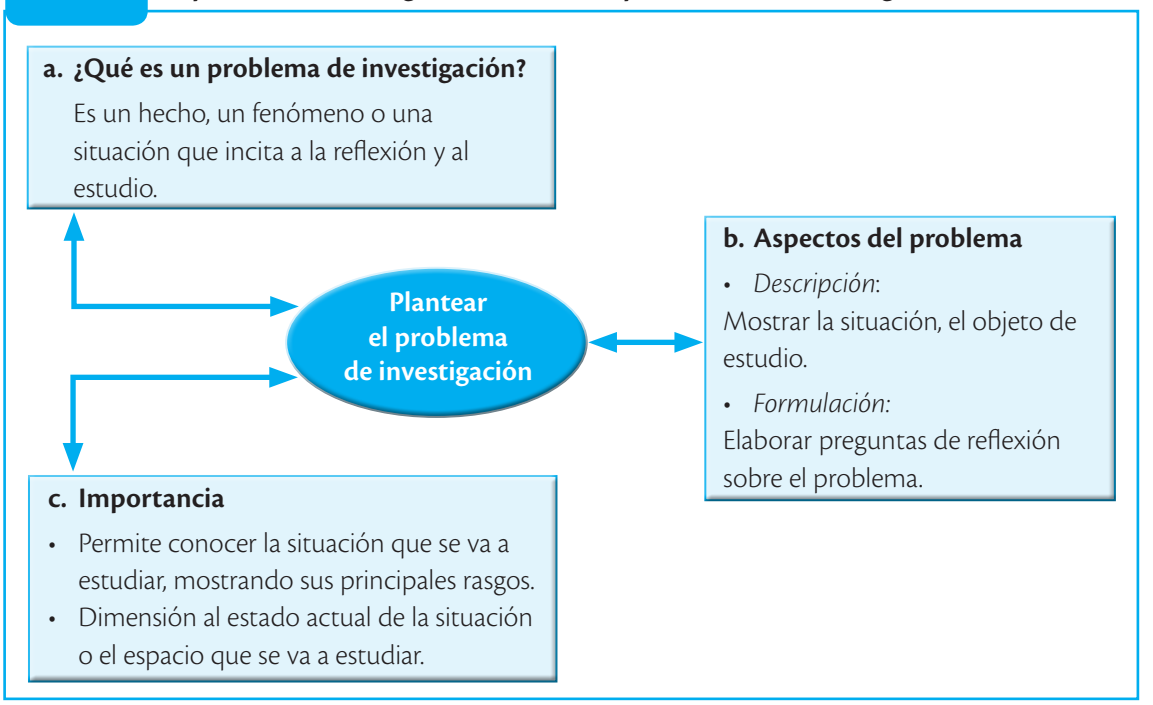
Las preguntas específicas están orientadas a interrogar sobre aspectos concretos del problema y no al problema en su totalidad, pero que en su conjunto conforman la totalidad (las preguntas específicas son subpreguntas de la pregunta general).

La Figura 7.4 ilustra los aspectos relevantes relacionados con el planteamiento del problema de investigación tales como el propio concepto de problema en investigación, los aspectos del problema y la importancia que representa una adecuada definición del problema en un proceso de investigación científica.

En los ejemplos 7.3 y 7.4 se muestran planteamientos de problemas de investigación en un proyecto de investigación:

www.FreeLibros.me

FIGURA 7.4 **El proceso de investigación. Plantear el problema de la investigación**



EJEMPLO 7.3 **Planteamiento del problema de investigación en el campo de la psicología y la educación**

Título de la investigación: *Compromiso pedagógico de la literatura infantil latinoamericana producida en la década de 1990 en la formación integral del niño y la niña.*

Planteamiento del problema

Enunciado del problema

Los niños viven en nuestros días en medio de una civilización tecnológica o postindustrial, rodeada cotidianamente por un flujo incesante y multiforme de estímulos y mensajes, así como inmersa en un perturbador y seductor universo de imágenes. Los medios de comunicación, en sus diversas expresiones y articulaciones, invaden, como protagonistas, cada aspecto y momento de la vida de las personas, en especial de la población infantil.

No extraña que en los países industrializados, como afirma Postman (1998), la formación del niño se deba en un 75% a la influencia de los medios de comunicación y sólo en un 25% a las instancias educativas tradicionales, mientras que a comienzos de la década de 1950 las proporciones eran inversas.

www.FreeLibros.me

Con la irrupción de la galaxia informática, telematizada, se anuncia, cargada de incógnitas y no carente de riesgos, una posterior revolución de la comunicación antropológica cultural, lógica-cognitiva y sensorioperceptiva, de proporciones inadvertidas (Nobile 1992) .

La presencia invasora de los medios de comunicación, complicada por la revolución telemática e informativa, operando junto a otros factores y diversas situaciones, propios de la actual civilización tecnológica, produce consecuencias claras en el desarrollo psicológico de la infancia y en su condición existencial. En el plano intelectual, hay que señalar la pasividad de recepción, el empobrecimiento del espíritu crítico y la inhibición de la autonomía de juicio, los cuales, unidos al empobrecimiento de la fantasía y la anulación de la creatividad, se traducen en un obstáculo para el desarrollo, la integración y el perfeccionamiento de las facultades superiores de la inteligencia y del pensamiento.

En el plano estrictamente intelectual, según diversos autores, la excesiva familiaridad del infante con la imagen no favorece el desarrollo cognitivo ni el de la fantasía; incluso, se perjudica el pensamiento creativo, determinando, a la vez, el progresivo deterioro de la curiosidad infantil.

La aparente fluidez expresiva lingüística, la riqueza léxica, el uso de un lenguaje más elaborado del niño televisivo no implican estructuras cognitivas más complejas, y son el resultado de adquisiciones superficiales y efímeras.

En tal contexto, una ponderada reflexión sobre la problemática sociopsicopedagógica de la civilización postindustrial, en especial en un deteriorado aspecto de la sociedad de consumo, hace recuperar y destacar el papel de la lectura crítica e inteligente, en la acepción bruneriana del término, en función antagónica y correctora de los nefastos efectos producidos por la invasora difusión de los medios audiovisuales, así como de la anunciada hegemonía de los nuevos medios electrónicos omnipresentes y multiperceptivos.

Si el libro es "un medio personal, capaz de poner en situación de objetiva autonomía cultural al usuario", la lectura, coloquio íntimo con el autor, reivindica su permanente función exquisitamente formativa como factor esencial de autorrealización de la persona, "actitud que implica toda la vida psíquica" (Guini, 1969), rica en reclamos interiores e interiorizados, principal instrumento de enculturización, proceso de reelaboración, interpretación y sistematización de la realidad, aproximación viva y vivificante a la lengua, que asume la irrenunciable función de ampliación de los conocimientos, así como de los horizontes intelectuales y culturales, de estructuración y potenciación de las facultades lógicas, de enriquecimiento del patrimonio lingüístico y de los medios expresivos.

Al ejercitarse en textos cualificados, en cuanto a lenguaje y contenidos, la lectura agudiza el espíritu crítico, refuerza la autonomía de juicio, educa el sentimiento estético, nutre la fantasía, ensancha la imaginación, habla a la afectividad, cultiva el sentimiento, descubre intereses más amplios y autónomos, contribuye a la promoción de una sólida conciencia moral y cívica, abierta a los ideales de comprensión humana y de solidaridad, por lo que resulta ideal para la formación integral de la persona, en especial cuando está en edad de formación.

www.FreeLibros.me

Al mismo tiempo, en el clamor ensordecedor de los medios de comunicación, la lectura permite el redescubrimiento de los valores psicopedagógicos del silencio y, algo aún más urgente, la recuperación de la dimensión de interioridad, favoreciendo el hábito de la reflexión y la introspección.

En una edad evolutiva, marcada por un flujo incesante de imágenes y acontecimientos, así como sumergida en un desestructurado y desestructurante torrente de información, que acaba por configurar todo de una forma confusa y fragmentaria, irreflexiva y pasiva, la lectura de calidad aporta los instrumentos, especialmente lógico-formales y lingüísticos, para su asimilación crítica e informada, lo mismo que para un uso activo de esos mismos medios, apareciendo como elemento de orden y disciplina interior.

En el siglo XXI , como supone Frabioni (1987), ante el niño y la niña, sentados en solitario delante de sus propios visos, pasarán miles de páginas de información memorizadas en gigantescos bancos de datos. Por otro lado, frente a un descenso general de las relaciones interpersonales, de convivencia y de posibilidades sociales de juegos de grupo se prevé el empobrecimiento y la fragmentación del saber, una invasión de imágenes y palabras escritas preelaboradas, con el consiguiente descenso de la creatividad y el desarrollo integral de la persona.

Las llamadas tecnologías de punta, especialmente la microelectrónica y la microinformática, tienden a influir y a acondicionar formas de razonamiento y pensamiento, dando mayor importancia al sistema informático y una comprensión unidimensional de la realidad.

El libro, depositario de un inestable patrimonio de conocimientos, cultura y civilización, lo mismo que síntesis de las creaciones de la fantasía y el ingenio humanos, es un medio fundamental privilegiado para cumplir una labor universal de educación, de entenderse como una transfusión de vida de una generación a otra.

Así como en la narrativa serena y optimista, lo mismo que rica en valores humanos, sociales y cívicos, además de interesante y atractiva, que plantea, a través de conocimientos culturales elegidos adecuadamente, ideales de vida positivos y modelos de conducta e identificación, indispensables en el proceso evolutivo, que ayuden en forma gradual al joven lector a tener contacto con la realidad y sus problemas, pueden ser un elemento importante en el desarrollo armónico e integral de la personalidad infantil, y asumir la urgente tarea de la regeneración espiritual de la juventud.

La literatura, especialmente la infantil, afirma Nobile (1992), debe alzarse en una sólida unidad de intenciones con la escuela, como último y generoso baluarte, contra la creciente marca del instinto, la ignorancia, la simultánea y destructora amenaza de un empobrecimiento humano y, sobre todo, el grave deterioro de la palabra como hecho espiritual, como descubrimiento personal y como fuente de desarrollo humano.

Cuando se habla aquí de literatura, se hace referencia a una literatura auténticamente formativa, capaz de ofrecer valores y contenidos culturales a una población infantil con un perfil cada vez más conformista, estandarizado y homogenizado, tanto en el lenguaje como en el vestido y en los comportamientos, lo mismo que llena de mitos consumistas por la persistencia del

www.FreeLibros.me

originario principio del placer, atento sólo a lo inmediato y egoísta, satisfacción de las propias e indisciplinadas pulsaciones del instinto.

En este sentido, la literatura es y debe ser fuente de expresión de humanidad y de racionalidad crítica, garantía de libertad de pensamiento, así como fuente inagotable de desarrollo humano integral; por ello y las razones expresadas que hoy enfrenta la población infantil, es importante estudiar el compromiso pedagógico de la literatura infantil en la formación de sus lectores.

Formulación del problema

En virtud de lo anterior, el presente trabajo se orienta al análisis formativo de la literatura infantil (latinoamericana), pero específicamente a la promoción de valores en los niños y niñas; por tanto, el estudio busca responder los siguientes interrogantes:

- ¿Qué se entiende por literatura infantil?
- ¿Existe compromiso pedagógico de la literatura infantil latinoamericana de la década de 1990 en la formación integral de sus lectores?
- ¿Hay relación entre la literatura infantil latinoamericana de la década del 90 y los valores que practican los niños y niñas que leen esta literatura?
- ¿La literatura infantil latinoamericana de la década del 90 ha sido un instrumento pedagógico de promoción de la dimensión social y psicoafectiva para el niño o la niña que lee esta literatura?

EJEMPLO 7.4 Planteamiento del problema de investigación aplicado al campo de los negocios

En el caso de una investigación sobre creatividad en las empresas nacionales, cuyo título podría ser *La industria manufacturera nacional frente a la innovación y la creatividad*, el planteamiento del problema sería el siguiente:

Planteamiento del problema

Enunciado del problema

En un mundo caracterizado por rápidos y complejos cambios, a los directivos empresariales cada vez se les plantean problemas más difíciles que no sólo afectan su organización, sino que influyen en el conjunto de la sociedad.

Por esto, cada vez tendrán mayor necesidad de enfrentar las situaciones de un modo creativo e innovador. Sin embargo, estudios recientes sobre el manejo de la creatividad y la

www.FreeLibros.me

innovación en las organizaciones empresariales revelan que en realidad el camino por recorrer es muy largo.

Investigadores como Sternberg y Lubar, y Kuczmarski (1997), en sus estudios sobre creatividad en las organizaciones, concluyen que la mayoría de los gerentes de alto nivel se atemorizan ante las personas creativas e innovadoras, pues las consideran incómodas y raras, y se niegan a abrazar lo novedoso y diferente, cuando desaniman y no premian el pensamiento innovador de sus empleados.

En general, pese a no reconocerlo, los directivos de las empresas nacionales entorpecen o no propician la creatividad, lo cual los pone en desventaja para asumir los retos del nuevo milenio.

La sociedad actual necesita directivos capaces de asumir un verdadero compromiso con sus trabajadores, más que desempeñar funciones preestablecidas, que sean innovadores y comprometidos con su propio desarrollo y el de la sociedad.

Formulación del problema

- ¿Tienen los empresarios nacionales de la industria manufacturera una actitud congruente entre lo que piensan y lo que hacen respecto a las personas creativas en sus empresas? (Formulación de una pregunta de investigación general).
- ¿Tienen los empresarios nacionales actitudes positivas frente a las personas creativas? (Pregunta específica).
- ¿Estimulan y promueven los empresarios nacionales la creatividad? (Pregunta específica).

Otro planteamiento de un problema de investigación podría ser el que se describe a continuación y que tiene que ver con los *Nuevos retos que enfrentan las empresas como consecuencia del nuevo paradigma mundial de la globalización y la competitividad.*

Planteamiento del problema

Enunciado del problema

Reconocidos investigadores en el ámbito de los negocios y de la economía, como Kenich Ohmae, Peter F. Drucker, Patricia Aburdene, John Naisbitt, Jack Nadel y Michael E. Porter, han analizado los cambios económicos del mundo, la naturaleza de los negocios y la necesidad de tener un nuevo paradigma, un nuevo enfoque y una nueva manera de pensar para interpretar y abordar la nueva realidad.

El planeta Tierra parece haberse convertido en una pequeña aldea. El avance de la globalización es una realidad; los individuos ya no son ciudadanos de una nación, sino ciudadanos del mundo; los sistemas de comunicación vía satélite, la telefonía celular, las autopistas de información vía Internet, la fibra óptica y los computadores están revolucionando los procesos de conocimiento.

www.FreeLibros.me

En el mundo de la política están rompiéndose los paradigmas básicos. Las estructuras de poder ligadas a grandes organizaciones sociales como partidos políticos y sindicatos dan paso ahora a los candidatos libres. Ya la fuerza no está en los partidos políticos, ahora está en las personas, en sus características humanas, individuales, personales.

Cambian los estilos de vida, se da una transformación total de los sistemas educativos del mundo. La calidad de la educación se caracterizará por un sistema de competencias individuales y por su infraestructura orientada hacia la investigación.

Se está en los albores de la era de la creatividad, de la era del conocimiento. Los países que tengan los mejores científicos, universidades, centros de investigación y centros de alta tecnología dominarán el mundo.

En el siglo XXI se dará un salto cualitativo en cuanto al aprovechamiento de las energías eólica y solar, y habrá mayor posibilidad de integración armónica con el entorno, y con formas energéticas menos agresivas para el medio biológico y para la biosfera.

Los negocios están enfrentando una paradoja, pues tienen oportunidades nunca vistas para aprovechar los nuevos mercados; sin embargo, éstos cambian de manera sustancial, reduciéndose o haciéndose intensamente competitivos. La apertura de los mercados mundiales generará una reestructuración en cada sector de los negocios, orientándolos a garantizar las siguientes estrategias: rapidez, bajo margen de ganancia por unidad, calidad en productos y servicios, diversidad de productos, descuentos, comodidad e innovación (Tucker, 1995).

Formulación del problema

- Frente a esta nueva realidad, ¿cómo deben prepararse las empresas nacionales y sus gerentes si quieren desenvolverse competitivamente? (Pregunta general).
- ¿Cómo favorece el gobierno la competitividad de las empresas nacionales?
- ¿Cuál es la actitud de los gremios ante los retos que impone el nuevo milenio?
- ¿Qué están haciendo las universidades para contribuir a la competitividad de las empresas?
- ¿Qué están haciendo los empresarios para responder al nuevo ambiente de competitividad?

RESUMEN

En investigación, problema es todo aquello que se convierte en objeto de reflexión y sobre el cual se percibe la necesidad de conocer y, por tanto, de estudiar. Es decir, problema no significa algo disfuncional, molesto o negativo, sino todo aquello que incite a ser conocido, y su conocimiento representa alguna utilidad ya sea para el investigador o para un tercero. De otra parte, en investigación no es suficiente visualizar un problema, es necesario plantearlo adecuadamente.

www.FreeLibros.me

Plantear el problema de investigación significa, entonces, enunciarlo y formularlo. La enunciación o descripción consiste en presentar el estado actual de la situación problema (la naturaleza y las dimensiones del mismo, los antecedentes, los hechos). Formular el problema consiste en plantear preguntas relevantes para responder con el desarrollo de la investigación a partir del estado de la situación que va a estudiarse. Las preguntas de investigación suelen plantearse en dos niveles: general (una pregunta) y específico (varias preguntas).

EJERCICIOS DE REPASO Y ANÁLISIS

1. ¿Qué se entiende por problema de investigación?
2. En investigación, ¿qué significa plantear el problema de investigación?
3. En el planteamiento del problema, ¿qué se entiende por enunciar el problema de investigación?
4. ¿En qué consiste la formulación de un problema de investigación?
5. ¿Cuál es la importancia de un buen planteamiento del problema en una propuesta de investigación?
6. Describa fuentes y orígenes de problemas de investigación en el campo de su disciplina, que puedan ser objeto de investigación.
7. En grupo, pregúntenles a cinco profesores cuáles serían algunos temas problémicos relevantes para investigar en su disciplina.
8. Realice una revisión bibliográfica que le permita conocer qué temas están investigando los docentes en su escuela o universidad y revise, si es posible, el documento que le permita conocer el planteamiento del problema de los casos identificados. Comente los hallazgos en clase.
9. Seleccione un tema de su interés y plantee el problema, luego preséntele este planteamiento a su profesor para que se lo revise.

www.FreeLibros.me

7.3 Objetivos de la investigación

Un aspecto definitivo en todo proceso de investigación es la definición de los objetivos o del rumbo que debe tomar la investigación que va a realizarse. Así, los objetivos son los propósitos del estudio, expresan el fin que pretende alcanzarse; por tanto, todo el desarrollo del trabajo de investigación se orientará a lograr estos objetivos.

Éstos deben ser claros y precisos para evitar confusiones o desviaciones; sin embargo, esto no implica que los objetivos no puedan modificarse durante la realización de la investigación, porque en algunos casos hay que hacerlo.

7.3.1 Cómo se redactan y definen los objetivos en la investigación

Puesto que todo objetivo implica la acción que se desea lograr, es importante tener en cuenta que al redactar los objetivos de la investigación deben utilizarse verbos en infinitivo.

No es necesario escribir preámbulos al momento de redactar los objetivos; incluso se recomienda expresar directamente el objetivo. Por ejemplo, si un objetivo es "hacer un análisis de la situación actual del sector de las artes gráficas en la ciudad de...", no es necesario agregar frases previas al objetivo como; "debido a que las empresas del sector de las artes gráficas atraviesan una situación económica difícil, en este estudio se pretende hacer un análisis...".

Otro aspecto muy importante en el momento de plantear los objetivos de la investigación es utilizar verbos que puedan lograrse o alcanzarse durante el desarrollo de la investigación:

Los verbos podrían ser:

Determinar	Verificar	Definir
Identificar	Diseñar	Conocer
Evaluar	Elaborar	Estudiar
Describir	Proponer	Plantear
Formular	Analizar	Corroborar

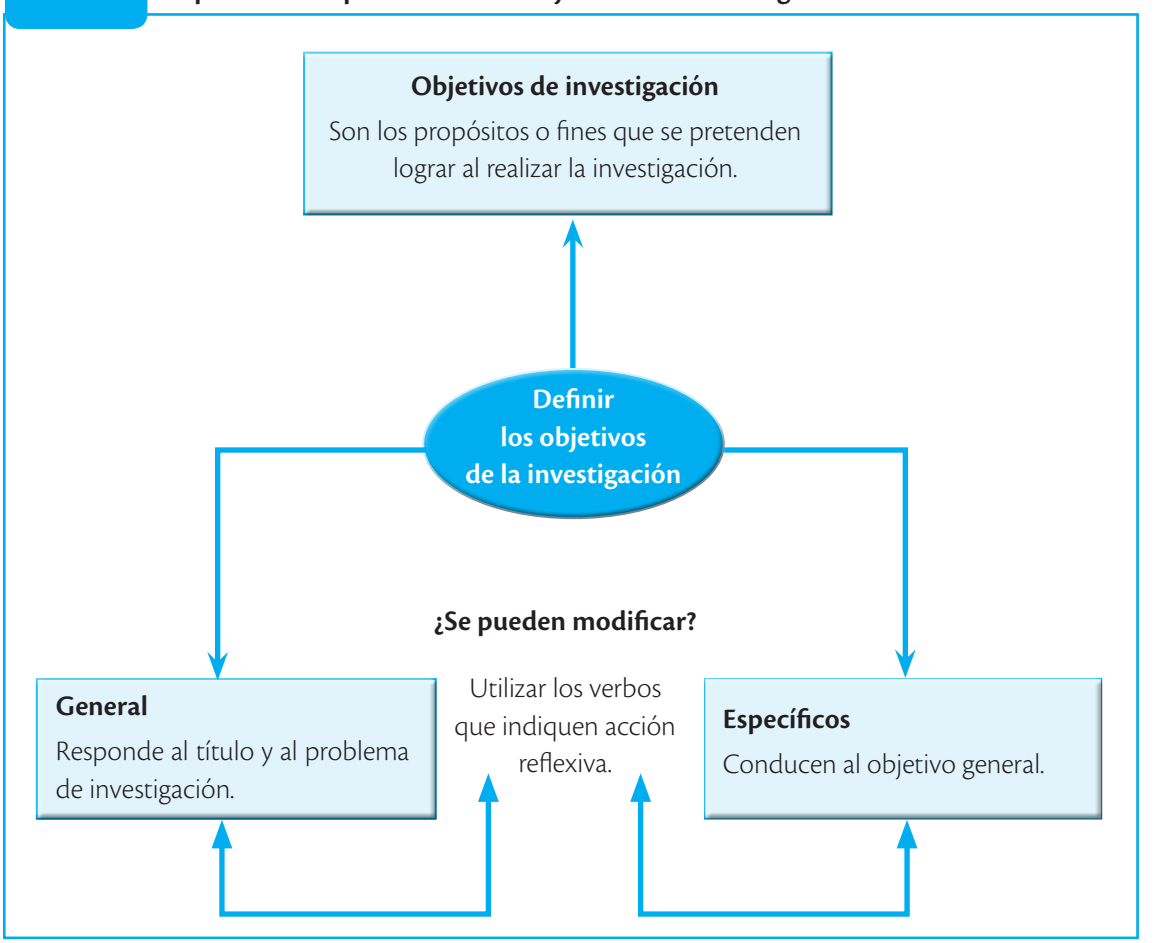
El uso de verbos como capacitar, cambiar, motivar, enseñar, mejorar y muchos otros que implican acciones finales debe ser prudente, porque estas acciones casi nunca se logran durante el progreso de la investigación, debido a que implican dedicarles tiempo y recursos y, muchas veces, tomar decisiones para desarrollar el objetivo propuesto.

www.FreeLibros.me

Sería deseable que todos los objetivos propuestos en un estudio pudieran llevarse a la práctica, con lo cual se contribuiría a solucionar problemas verdaderos. Sin embargo, el hecho de que no se apliquen no debe ser motivo de frustraciones, porque en el pregrado el solo hecho de realizar el ejercicio teórico de la investigación es ya un gran avance; y en el caso de la maestría, lo importante es la reflexión académica y fundamentada que pueda hacerse sobre el quehacer investigativo y la actitud que al respecto se cree.

La Figura 7.5 muestra los aspectos clave para la definición de los objetivos de la investigación, comenzando por un claro concepto de estos y la categorización en sus niveles de general y específicos, así como el tipo de verbos que han de utilizarse para su definición y la posibilidad de modificarse justificadamente.

FIGURA 7.5 **Aspectos clave para definir los objetivos de la investigación**



En los ejemplos 7.5 y 7.6 se muestran las limitaciones para alcanzar algunos objetivos.

www.FreeLibros.me

EJEMPLO 7.5

Si un estudiante en uno de sus objetivos se propone "Mejorar las condiciones ambientales de las empresas para generar un mejor entorno laboral", este objetivo implica realizar un diagnóstico de las condiciones ambientales actuales en las empresas; identificar dónde es necesario mejorar, proponer las estrategias para el mejoramiento de las condiciones ambientales y poner en marcha tales estrategias hasta producir el cambio o la mejora.

Esta última parte, que considera implementar las estrategias, casi nunca la aplica el estudiante; él sólo busca proponer estrategias para mejorar y hasta ahí llega su trabajo.

EJEMPLO 7.6

Al plantearse el objetivo "Disminuir el nivel de estrés en los miembros de familias y de sus miembros en enfermedad terminal", este es poco realizable durante el desarrollo de un trabajo de grado. Posiblemente el estudiante logre plantear estrategias para disminuir el grado de estrés, pero ello no significa que en realidad disminuya.

Para ello se requiere desarrollar un programa que implica tiempo y recursos, y el estudiante usualmente no llega a esa etapa en su trabajo de grado.

7.3.2 Objetivos: general y específicos

En toda investigación se requiere plantear dos niveles en los objetivos: El general y los específicos.

El objetivo general Debe reflejar la esencia del planteamiento del problema y la idea expresada en el título del proyecto de investigación.

Los objetivos específicos Se desprenden del general y deben formularse de forma que estén orientados al logro del objetivo general, es decir, que cada objetivo específico esté diseñado para lograr un aspecto de aquél; y todos en su conjunto, la totalidad del objetivo general. Los objetivos específicos son los pasos que se dan para lograr el objetivo general. Los ejemplos 7.7 y 7.8 muestran objetivos de investigación.

www.FreeLibros.me

EJEMPLO 7.7 **Objetivos en el campo de la economía y los negocios**

Título de la investigación: *Actitud de los directivos nacionales respecto a la innovación y la creatividad en sus empresas.*

Objetivos de la investigación:

Objetivo general

Analizar la actitud de los directivos nacionales de la industria manufacturera respecto a la innovación y creatividad en sus empresas.

Objetivos específicos

1. Elaborar un marco teórico sobre la actitud de los directivos de empresas respecto a la innovación y creatividad en las organizaciones productivas.
2. Diseñar un instrumento de medición de aptitudes con sus respectivas pruebas de validez y confiabilidad para aplicarlo a la población seleccionada.
3. Identificar la actitud de los gerentes de las 100 empresas más grandes del país respecto a la innovación y la creatividad en ellas.
4. Analizar los resultados obtenidos según subsector y demás variables sociodemográficas de las empresas estudiadas.

Título de la investigación: *Capacidad de adaptación de las empresas a los nuevos retos de competitividad y globalización.*

Objetivos de la investigación:

Objetivo general

Analizar la forma en que están preparándose las empresas nacionales para afrontar los retos de la competitividad y los demás cambios que caracterizan el nuevo ambiente de los negocios.

Objetivos específicos

1. Describir los retos que deben afrontar los negocios en el paradigma de la competitividad y la globalización.
2. Identificar las acciones que está realizando el gobierno para apoyar a las organizaciones ante los nuevos retos de la competitividad.
3. Conocer las acciones que han emprendido las organizaciones para hacerlas competitivas en un escenario combatiente y globalizado.
4. Evaluar qué papel desempeñan las instituciones universitarias en el proceso de preparación de organizaciones empresariales, para afrontar los retos de la nueva teoría económica y administrativa.

www.FreeLibros.me

Título de la investigación: *Impacto de las medidas económicas del gobierno en la actividad económica del país.*

Objetivos de la investigación:

Objetivo general

Evaluar el impacto de las medidas económicas del gobierno actual en la actividad empresarial.

Objetivos específicos

1. Describir las medidas económicas adoptadas por el gobierno en lo que va de su mandato.
2. Analizar el comportamiento de las tasas de interés, de la inflación y de las tasas de cambio durante el gobierno actual.
3. Identificar la opinión de los empresarios respecto a las medidas económicas del gobierno actual.
4. Identificar la opinión de los trabajadores sobre las medidas económicas del gobierno actual.
5. Analizar indicadores como el desempleo y la inversión social, y confrontarlas con la política económica del gobierno actual.

Título de la investigación: *Análisis comparativo entre el sistema de costos basado en actividades y el sistema de costeo tradicional.*

Objetivos de la investigación:

Objetivo general

Hacer un análisis comparativo entre el sistema de costeo basado en actividades y el sistema de costeo tradicional.

Objetivos específicos

1. Mostrar los fundamentos básicos del sistema de costeo basado en actividades y su diferencia con el sistema de costeo tradicional.
2. Evaluar las ventajas del sistema de costeo basado en actividades respecto al sistema de costeo tradicional.
3. Mostrar la importancia del costeo basado en actividades como herramienta de apoyo para la toma de decisiones gerenciales.

www.FreeLibros.me

EJEMPLO 7.8 **Objetivos de investigación en temas de psicología y educación**

Título de la investigación: *Análisis de la relación entre el tipo de valores practicados por los estudiantes universitarios y su rendimiento académico.*

Objetivos de la investigación:

Objetivo general

Analizar la relación entre el tipo de valores practicados por los estudiantes universitarios y su rendimiento académico.

Objetivos específicos

1. Identificar el tipo de valores practicados por los estudiantes universitarios.
2. Analizar el tipo de valores que practican los estudiantes universitarios de más alto rendimiento.
3. Conocer el tipo de valores que practican los estudiantes universitarios de bajo rendimiento.
4. Analizar la diferencia significativa entre los valores que practican los estudiantes universitarios de más alto rendimiento y los que practican los estudiantes de bajo rendimiento.
5. Evaluar la relación entre el tipo de valores practicados por los estudiantes objeto de estudio y el rendimiento académico de los mismos.

Título de la investigación: *Estudio sobre el estrés en la familia urbana en Calamba.*

Objetivos de la investigación:

Objetivo general

Analizar los factores generadores de estrés en la familia urbana calambana y las implicaciones que estos factores generan en el funcionamiento familiar.

Objetivos específicos

1. Identificar los factores (internos y del entorno) generadores de estrés en las familias calambanas.
2. Conocer las implicaciones del estrés en las familias estudiadas y en cada uno de sus miembros.
3. Analizar las estrategias que adoptan las familias calambanas para manejar el estrés.
4. Evaluar diferencias significativas entre los factores generadores de estrés y las estrategias utilizadas para el manejo de éste en las familias según estratos y nivel académico de sus miembros.

www.FreeLibros.me

Título de la investigación: *Estudio exploratorio sobre creatividad verbal, producción y comprensión de metáforas en niños de 4 a 12 años de edad.*

Objetivos de la investigación:

- ***Objetivo general***

Analizar la capacidad verbal, tanto para la producción como para la comprensión de metáforas, de niños y niñas de 4 a 12 años de edad, de estrato socioeconómico medio de la ciudad *XMJZO*.

- ***Objetivos específicos***

1. Evaluar la capacidad de producción de diferentes tipos de metáforas, en una muestra de niños y niñas de la ciudad *XMJZO*.
2. Evaluar la capacidad de comprensión de diferentes tipos de metáforas en la muestra de niños y niñas de la ciudad *XMJZO*.
3. Identificar la relación entre la capacidad de producción y de comprensión de diferentes tipos de metáforas en la muestra de niños y niñas que se van a evaluar.
4. Analizar diferencias y similitudes entre la capacidad de producción y comprensión de los diferentes tipos de metáforas, por parte de la muestra de niños y niñas de la ciudad *XMJZO*.
5. Discutir los resultados del trabajo de campo con los resultados de investigaciones realizadas por H. Gardner, citadas en el marco de referencia de este estudio.

Título de la investigación: *Evaluación de los efectos de un programa para la prevención de comportamientos agresivos de jóvenes en espectáculos masivos.*

Objetivos de la investigación:

- ***Objetivo general***

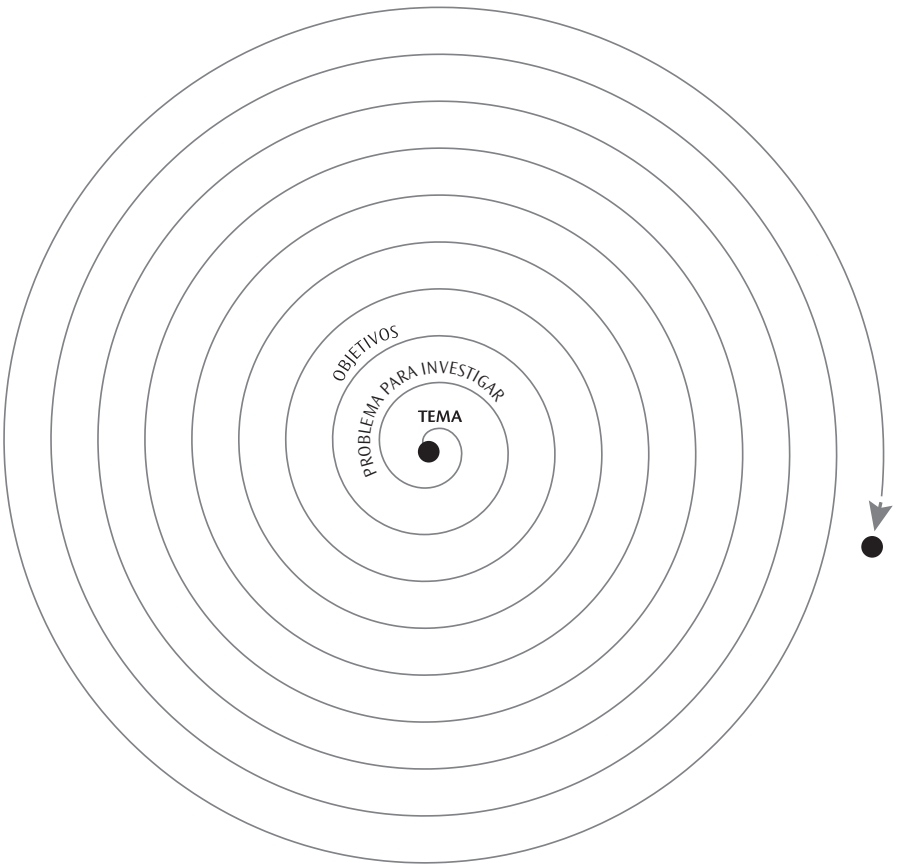
Evaluar los efectos de un programa de prevención de comportamientos agresivos de jóvenes que asisten a espectáculos masivos.

- ***Objetivos específicos***

1. Diseñar o adaptar un instrumento de evaluación, que sea válido para la medición del impacto del programa de prevención de comportamientos agresivos de parte de los jóvenes que asisten a espectáculos masivos.
2. Hacer el trabajo de evaluación de impacto.
3. Analizar la información obtenida en el trabajo de campo, orientada a evaluar el impacto del programa de prevención de los comportamientos agresivos de los jóvenes que asisten a espectáculos masivos.
4. Formular recomendaciones sobre los efectos del programa, así como evaluar y diseñar estrategias para la difusión del estudio, si los resultados de la evaluación son positivos.

www.FreeLibros.me

FIGURA 7.6 **El proceso de investigación. Establecimiento de los objetivos de la investigación**



RESUMEN

En el proceso de investigación, los objetivos definen el rumbo que debe tomar la investigación que va a realizarse; son los propósitos del estudio, y expresan el fin que pretende alcanzarse; el desarrollo del trabajo de investigación se orientará a lograr estos objetivos. No obstante, estos objetivos pueden modificarse durante la realización de la investigación, siempre que haya razones para hacerlo.

Los objetivos en investigación deben ser claros, precisos y realizables, y su redacción exige la utilización de verbos en infinitivo que implican acción reflexiva más que operativa. Estos objetivos se han de plantear en dos niveles: general y específico. El objetivo general debe reflejar la esencia del planteamiento del problema y la idea expresada en el título del proyecto de investigación. Los objetivos específicos se desprenden del general y deben formularse orientados al logro de éste, es decir, cada objetivo específico debe diseñarse para lograr un aspecto de aquél; y todos en su conjunto, la totalidad del objetivo general.

www.FreeLibros.me

EJERCICIOS DE REPASO Y ANÁLISIS

1. ¿Por qué es importante establecer objetivos en una investigación?
2. ¿Cuáles son los dos niveles de objetivos que se plantean en toda investigación?
3. Revise los objetivos de alguna investigación ya publicada en el campo de su disciplina y socialícelos en la clase.
4. Seleccione un tema de investigación y haga un breve planteamiento del problema. Además, establezca objetivos de investigación para el tema respectivo.
5. Suponga que usted ha sido elegido para dar un concepto sobre la siguiente propuesta de investigación.

Título

Estrategias para el mejoramiento de la calidad del servicio en el sector hotelero nacional de cuatro y cinco estrellas.

Objetivos de la investigación

• *Objetivo general*

Elaborar un diagnóstico organizacional sobre la calidad del servicio al cliente en los hoteles de cuatro y cinco estrellas existentes en el país.

• *Objetivos específicos*

1. Conocer las fortalezas y debilidades de los hoteles objeto del estudio.
2. Conocer la percepción del cliente respecto al servicio de tales hoteles.
3. Capacitar a los trabajadores de los hoteles objeto del estudio en técnicas de servicio.
4. Promover el programa de servicio.

Usted debe conceptuar respecto a:

a. La pertinencia del tema en el campo de las ciencias económicas y administrativas.
b. La claridad del título de la investigación.
c. La correspondencia entre el título de la investigación y los objetivos de la misma.
d. La correspondencia entre el objetivo general y los objetivos específicos.
e. El grado de realismo para el logro o alcance de los objetivos.

En caso de tener observaciones en cualquiera de los criterios anteriores, haga las recomendaciones necesarias. Suponiendo que usted es el asesor del estudio, haga las correcciones conservando el título original.

6. Suponga que su tema de investigación es: "Importancia de los liderazgos masculino y femenino en la gerencia de las empresas del sector financiero del país". Plantee los objetivos general y específicos para este tema.
7. Suponga que su tema de investigación es: "Análisis de los principales factores desencadenantes de conductas delictivas en los adolescentes". En este caso plantee:
   a. El problema de investigación.
   b. Los objetivos general y específicos del estudio.

www.FreeLibros.me

7.4 Justificación y delimitación de la investigación

Toda investigación está orientada a la resolución de algún problema; por consiguiente, es necesario *justificar,* o exponer, los motivos que merecen la investigación. Asimismo, debe determinarse su cubrimiento o dimensión para conocer su viabilidad.

7.4.1 Criterios de justificación

De acuerdo con Méndez, la justificación de una investigación puede ser de carácter teórico, práctico o metodológico (Méndez, 1995).

Justificación teórica

En investigación hay una *justificación teórica* cuando el propósito del estudio es generar reflexión y debate académico sobre el conocimiento existente, confrontar una teoría, contrastar resultados o hacer epistemología del conocimiento existente.

En el caso de las ciencias económico-administrativas, un trabajo investigativo tiene justificación teórica cuando se cuestiona una teoría administrativa o una económica (es decir, los principios que la soportan), su proceso de implantación o sus resultados.

Cuando en una investigación se busca mostrar las soluciones de un modelo, está haciéndose una justificación teórica.

Si en una investigación se proponen nuevos paradigmas o se hace una reflexión epistemológica, se tiene una justificación eminentemente teórica, aunque al implementarla se vuelve práctica, ya que, como afirma López Cerezo (1988), toda investigación en alguna medida tiene la doble implicación, teórica y práctica.

La justificación teórica es la base de los programas de doctorado y algunos programas de maestría donde se tiene como objetivo la reflexión académica.

Justificación práctica

Se considera que una investigación tiene *justificación práctica* cuando su desarrollo ayuda a resolver un problema o, por lo menos, propone estrategias que al aplicarse contribuirían a resolverlo.

Los estudios de investigación de pregrado y de posgrado, en el campo de las ciencias económicas y administrativas, en general son de carácter práctico, o bien, describen o analizan un problema o plantean estrategias que podrían solucionar problemas reales si se llevaran a cabo.

Cuando en un trabajo de grado se realiza un análisis económico de un sector de la producción, su justificación es práctica porque genera información que podría utilizarse para tomar medidas tendientes a mejorar ese sector.

www.FreeLibros.me

Cuando un trabajo de grado se orienta a conocer los factores de motivación más utilizados en un determinado sector económico o en una empresa, su justificación es práctica, porque, al igual que en el caso del análisis del sector, la información sirve para actuar sobre la empresa, para mejorar o realizar *benchmarking* en otras organizaciones, o para confrontar la teoría sobre el tema.

Justificación metodológica

En investigación científica, la *justificación metodológica* del estudio se da cuando el proyecto que se va a realizar propone un nuevo método o una nueva estrategia para generar conocimiento válido y confiable.

7.4.2 Limitaciones del estudio o de la investigación

Una vez justificada la investigación, es necesario plantear las limitaciones dentro de las cuales se realizará (no todos los estudios tienen las mismas limitaciones, pues cada investigación es particular).

Las limitaciones en un proyecto de investigación pueden referirse a:

Limitaciones de tiempo Es necesario determinar cuál será el período, sea retrospectivo o prospectivo, dentro del cual se realizará el estudio del hecho, la situación, el fenómeno o población investigados. Por ejemplo, si el interés es estudiar el comportamiento de un sector económico, o los indicadores económicos de un país, es indispensable definir durante qué período (años, meses) se realizará tal análisis.

Limitaciones de espacio o territorio Son aquellas demarcaciones del espacio geográfico dentro del cual tendrá lugar una investigación. Las investigaciones pueden limitarse a una zona de una ciudad, a una ciudad, una región, un país, un continente, etcétera.

Limitaciones de recursos Mencionan la disponibilidad de recursos financieros para la realización del proyecto de investigación.

Además de estas limitaciones, en un proyecto de investigación puede haber limitaciones de información, población disponible para el estudio, dificultad de acceso a la misma, etcétera.

El investigador debe explicitar las limitaciones del proyecto con el propósito de facilitar su viabilidad.

EJEMPLO 7.9 **Enunciados de limitaciones de tiempo**

- Diagnóstico del sector de las flores en los últimos 12 años (1997-2009).
- Evaluación de la tasa de desempleo durante los últimos cuatro gobiernos.
- Análisis de las estrategias didácticas utilizadas en el proceso enseñanza-aprendizaje de la educación superior en los últimos 10 años.

www.FreeLibros.me

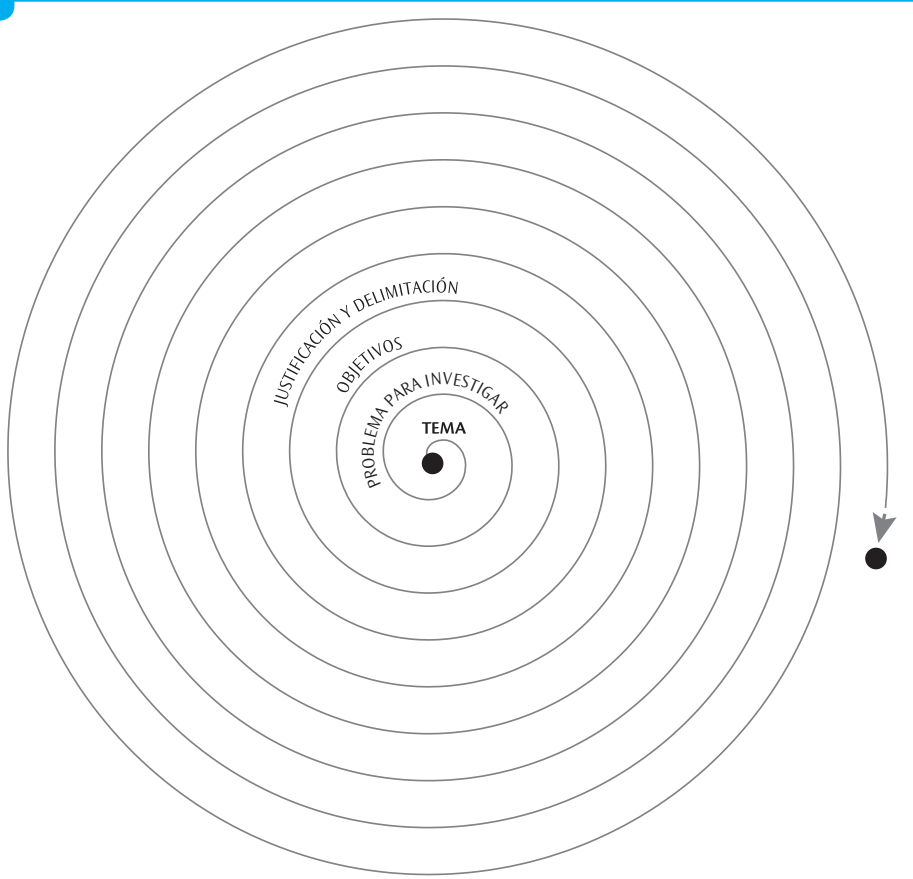
Existen investigaciones que se circunscriben a una determinada región, y eso se conoce como limitaciones geográficas.

EJEMPLO 7.10 **Enunciados de limitaciones de espacio o territorio**

- Cuando se dice que el estudio se realizará en empresas localizadas en la ciudad *PQTL* o en las empresas latinoamericanas, se hace referencia a una limitación geográfica.
- Afirmar que un estudio se realizará en una determinada ciudad o zona geográfica implica delimitar este estudio geográfica o territorialmente.

En general, puede hablarse de varias limitaciones según las características de la población o muestra sobre la cual recae el estudio. Estas características se detallarán cuando se haga referencia a la población y muestra objeto de estudio.

FIGURA 7.7 **El proceso de investigación. Justificación y delimitación de la investigación**



www.FreeLibros.me

La Figura 7.7 sintetiza la justificación como las razones de carácter práctico, teórico y metodológico que se tienen para realizar el estudio, así como sus delimitaciones en los ámbitos espacial, cronológico y/o sociodemográfico.

RESUMEN

En investigación, la justificación se refiere a las razones del porqué y el para qué de la investigación que se va a realizar, es decir, justificar una investigación consiste en exponer los motivos por los cuales es importante llevar a cabo el respectivo estudio. Al respecto, suele haber tres dimensiones o tipos de justificación: teórica, práctica y metodológica. Algunas investigaciones pueden requerir los tres tipos de justificación, otras dos tipos y otras solo uno. Ello depende de las particularidades de cada investigación.

La delimitación o el alcance en investigación se refiere a la dimensión o al cubrimiento que ésta tendrá en el espacio geográfico, período de tiempo y perfil sociodemográfico del objeto de estudio.

EJERCICIOS DE REPASO Y ANÁLISIS

1. ¿Qué significa justificar y delimitar una investigación?
2. Explique las distintas formas de justificar una investigación.
3. Seleccione un tema de investigación y justifique por qué debe realizarse el estudio.
4. Revise los ejemplos de los temas propuestos en el tema relacionado con la formulación de objetivos que justifiquen la realización de esos estudios.
5. Suponga que quiere realizar un estudio sobre el concepto de ser humano que han manejado las distintas teorías administrativas a través del tiempo. ¿Cuál es la justificación más adecuada para este estudio? Justifique su respuesta.
6. Suponga que recibe una propuesta de investigación cuyo título es: "Nuevo modelo para abordar el estudio de las organizaciones empresariales". El autor del estudio dice que su trabajo se justifica porque es un nuevo esquema de analizar las organizaciones y, por tanto, es práctico. ¿Qué opina de este argumento para justificar el estudio?

www.FreeLibros.me

7.5 Tipos de investigación

En la ciencia existen diferentes tipos de investigación y es necesario conocer sus características para saber cuál de ellos se ajusta mejor a la investigación que va a realizarse.

Aunque no hay acuerdo entre los distintos tratadistas en torno a la clasificación de los tipos de investigación, en este libro los principales tipos de investigación son los siguientes:

- Histórica
- Documental
- Descriptiva
- Correlacional
- Explicativa o causal
- Estudio de caso
- Experimental
- Otros

La elección o selección del tipo de investigación depende, en alto grado, del objetivo del estudio del problema de investigación y de las hipótesis que se formulen en el trabajo que se va a realizar, así como de la concepción epistemológica y filosófica de la persona o del equipo investigador.

7.5.1 Investigación histórica

En general, y siguiendo a Salkind (1998), "la investigación histórica se orienta a estudiar los sucesos del pasado. Analiza la relación de esos sucesos con otros eventos de la época y con sucesos presentes" (p. 12). En síntesis, se busca entender el pasado y su relación con el presente y el futuro.

Para Cerda (1998), la investigación histórica significa estudiar y examinar los fenómenos, como producto de un determinado desarrollo, desde el punto de vista como han aparecido, evolucionado y llegado al estado actual. El autor aclara que la palabra historia tiene generalmente dos significados, sobre los cuales se debe tener claridad cuando se trata de investigación. Uno hace referencia a las acciones humanas en el tiempo y sus correspondientes estudio y registro. Tal es la denominada historia-conocimiento, la de los libros y la de la historiografía,

www.FreeLibros.me

es decir, una historia definida, organizada y traducida en términos de unidades cronológicas y temporales.

El otro significado hace referencia al devenir de las sociedades y de los acontecimientos que se generan en el tiempo y en el espacio, los cuales el autor denomina "historia real", que es el objeto de estudio del investigador histórico.

Además, es necesario aclarar que la investigación histórica se aplica no sólo a la historia como disciplina, sino a toda área del conocimiento que quiera analizar los hechos del pasado, así como relacionarlos con hechos del presente y supuestos del futuro.

La investigación histórica tiene como principales fuentes y técnicas de obtención de la información, la revisión documental, los vestigios y objetos reales, así como en algunos casos personas que tuvieron relación directa con los hechos estudiados o quienes, aunque no tuvieron relación directa con esos hechos, cuentan con información válida y confiable sobre los mismos.

EJEMPLO 7.11 **Algunos títulos de investigación histórica**

En economía y administración

- Análisis de la Revolución Industrial y sus efectos sobre el desarrollo actual.
- Influencia de las experiencias administrativas de principios del siglo XX en las modernas teorías administrativas.

En psicología y educación

- Evolución de los principales paradigmas o enfoques en la psicología.
- Análisis histórico de estudios de psicopatología de la comunicación.
- Evolución de las estrategias pedagógicas en la educación básica primaria.
- Orígenes y evolución de los estudios sobre inteligencia, así como su relación con las estrategias pedagógicas.

7.5.2 Investigación documental

La investigación documental consiste en un análisis de la información escrita sobre un determinado tema, con el propósito de establecer relaciones, diferencias, etapas, posturas o estado actual del conocimiento respecto al tema objeto de estudio.

De acuerdo con Casares Hernández, et al. (1995), la investigación documental depende fundamentalmente de la información que se obtiene o se consulta en documentos, entendiendo por éstos todo material al que se puede acudir como fuente de referencia, sin que se altere su naturaleza o sentido, los cuales aportan información o dan testimonio de una realidad o un acontecimiento.

Para los autores mencionados, las principales fuentes documentales son: documentos escritos (libros, periódicos, revistas, actas notariales, tratados, conferencias escritas,

www.FreeLibros.me

etcétera), documentos fílmicos (películas, diapositivas, etcétera) y documentos grabados (discos, cintas, casetes, disquetes, etcétera).

En la investigación documental es importante mencionar las investigaciones denominadas "estado del arte", las cuales se caracterizan por abordar problemas de carácter teórico y empírico y que son relevantes en el tema objeto de estudio.

Los "estados del arte" son estudios cuyo propósito es mostrar el estado actual del conocimiento en un determinado campo o de un tema específico. En consecuencia, tales estudios muestran el conocimiento relevante y actualizado, las tendencias, los núcleos problemáticos, los vacíos, los principales enfoques o escuelas, las coincidencias y las diferencias entre esas hipótesis y los avances sobre un tema determinado. Es importante aclarar que los estados de arte no son un inventario del conocimiento de un tema objeto de estudio, ya que implican análisis de la información documental revisada, tomando en cuenta consideraciones epistemológicas y criterios contextualizadores en los que se dieron y se dan estos conocimientos.

Es importante no confundir "estado del arte" con marco teórico. El "estado del arte", como ya se mencionó, es un tipo de investigación documental, mientras que el marco teórico, que también es documental, es un aspecto constitutivo de toda investigación científica, que tiene como función básica servir de fundamento teórico de las investigaciones científicas. (Para mayor información sobre marco teórico véase la sección 7.6 Marco de referencia de la investigación).

EJEMPLO 7.12 **Enunciados de investigación documental**

En ciencias económico-administrativas

- Estudios relacionados sobre teoría de las organizaciones
- Estudios relacionados sobre modelos y doctrinas económicas.

En educación y psicología

- Estudios sobre teorías del aprendizaje.
- Estudios sobre enfoques y paradigmas psicológicos.

EJEMPLO 7.13 **Enunciados de investigación sobre estado del arte**

En ciencias económico-administrativas

- Estado del arte sobre cultura organizacional en América Latina.
- Estado del arte sobre competitividad empresarial en los países en desarrollo.

En educación y psicología

- Estado del arte sobre innovaciones pedagógicas en la educación superior.
- Estado del arte de la enseñanza de la psicología clínica.
- Estado del arte sobre la psicoterapia familiar.

www.FreeLibros.me

7.5.3 Investigación descriptiva

Se considera como investigación descriptiva aquella en que, como afirma Salkind (1998), "se reseñan las características o rasgos de la situación o fenómeno objeto de estudio" (p. 11).

Según Cerda (1998), "tradicionalmente se define la palabra *describir* como el acto de representar, reproducir o figurar a personas, animales o cosas…"; y agrega: "Se deben describir aquellos aspectos más característicos, distintivos y particulares de estas personas, situaciones o cosas, o sea, aquellas propiedades que las hacen reconocibles a los ojos de los demás" (p. 71).

De acuerdo con este autor, una de las funciones principales de la investigación descriptiva es la capacidad para seleccionar las características fundamentales del objeto de estudio y su descripción detallada de las partes, categorías o clases de ese objeto.

La investigación descriptiva es uno de los tipos o procedimientos investigativos más populares y utilizados por los principiantes en la actividad investigativa. Los trabajos de grado, en los pregrados y en muchas de las maestrías, son estudios de carácter eminentemente descriptivo. En tales estudios se muestran, narran, reseñan o identifican hechos, situaciones, rasgos, características de un objeto de estudio, o se diseñan productos, modelos, prototipos, guías, etcétera, pero no se dan explicaciones o razones de las situaciones, los hechos, los fenómenos, etcétera.

Para muchos expertos, la investigación descriptiva es un nivel básico de investigación, el cual se convierte en la base de otros tipos de investigación; además, agregan que la mayoría de los tipos de estudios tienen, de una u otra forma, aspectos de carácter descriptivo.

Esta investigación se guía por las preguntas de investigación que se formula el investigador; cuando se plantean hipótesis en los estudios descriptivos, éstas se formulan a nivel descriptivo y se prueban esas hipótesis.

La investigación descriptiva se soporta principalmente en técnicas como la encuesta, la entrevista, la observación y la revisión documental.

EJEMPLO 7.14 **Algunos temas de investigación descriptiva**

- Estudios de carácter diagnóstico.
- Diseños de guías, modelos, productos, prototipos, etcétera.
- Estudios de mercado.
- Estudios orientados a la descripción o identificación de rasgos o características de un objeto de estudio.
- Estudios de descripción de conductas, de actitudes, de perfiles, etcétera.
- Estudios de tiempos y movimientos.

www.FreeLibros.me

7.5.4 Investigación correlacional

Para Salkind (1998), la *investigación correlacional* tiene como propósito mostrar o examinar la relación entre variables o resultados de variables. De acuerdo con este autor, uno de los puntos importantes respecto a la investigación correlacional es examinar relaciones entre variables o sus resultados, pero en ningún momento explica que una sea la causa de la otra. En otras palabras, la correlación examina asociaciones pero no relaciones causales, donde un cambio en un factor influye directamente en un cambio en otro.

EJEMPLO 7.15 **Algunos títulos de investigaciones correlacionales**

En el campo económico y empresarial

- Las investigaciones orientadas a medir el efecto de las políticas de una empresa en el rendimiento de los trabajadores.
- Las investigaciones cuyos propósitos son conocer la relación entre el estado motivacional de los trabajadores y el nivel de productividad de los mismos.
- Las investigaciones orientadas a medir el impacto de un nuevo modelo administrativo en la actitud de los directivos empresariales.

En el campo de la psicología y la educación

- Estudios orientados a evaluar la relación entre el rendimiento académico de los hijos y el tipo de unión familiar de los padres de familia.
- Estudios orientados a evaluar o medir el grado de relación entre estrategias pedagógicas y el nivel de aprendizaje.
- Estudios que buscan conocer la relación entre tipo de prueba de evaluación y procesos y técnicas de aprendizaje.
- Estudios para identificar la relación entre los tipos de valores practicados en el ambiente académico institucional y el estado emocional de la comunidad estudiantil.
- Estudios para identificar la relación entre ciclos de vida marital y el estado emocional de los diferentes miembros de la familia.
- Estudios encaminados a evaluar la relación entre los estilos de comunicación intrafamiliar y el rendimiento académico de los hijos.
- Estudios orientados a evaluar la relación entre conductas suicidas y los entornos familiar y escolar de los adolescentes.
- Estudios orientados a evaluar la relación entre hábitos de consumo de drogas psicotrópicas y el ambiente social de los adolescentes.
- Estudios orientados a evaluar la relación entre causas de separación de las parejas y el entorno social de los miembros de la pareja.

www.FreeLibros.me

7.5.5 Investigación explicativa o causal

Así como se afirma que la investigación descriptiva es el nivel básico de la investigación científica, la investigación explicativa o causal es para muchos expertos el ideal y nivel culmen de la investigación no experimental, el modelo de investigación "no experimental" por antonomasia. La investigación explicativa tiene como fundamento la prueba de hipótesis y busca que las conclusiones lleven a la formulación o al contraste de leyes o principios científicos.

Las investigaciones en que el investigador se plantea como objetivos estudiar el porqué de las cosas, los hechos, los fenómenos o las situaciones, se denominan explicativas.

En la investigación explicativa se analizan causas y efectos de la relación entre variables.

EJEMPLO 7.16 **Algunos títulos de investigaciones explicativas o causales**

En el campo económico y empresarial

- Conocer las principales causas que llevan a la quiebra a las pequeñas y medianas empresas.
- Analizar los factores de competitividad de las empresas de un determinado país o sector económico.
- Analizar las estrategias o los factores de éxito o el fracaso de una empresa o un sector económico.

En psicología y educación

- Estudios orientados a explicar los efectos y las implicaciones de la clonación en las relaciones de pareja.
- Estudios orientados a explicar las principales razones de suicidio de adolescentes en los países subdesarrollados.
- Estudios orientados a explicar los efectos del consumo de drogas psicotrópicas por los adolescentes, según sus condiciones mentales y emocionales.
- Estudios que buscan explicar los efectos del aborto en la vida emocional de las mujeres que abortan.
- Estudios orientados a explicar las relaciones entre ambiente académico y rendimiento escolar.
- Estudios orientados a explicar el nuevo papel del docente en la sociedad del conocimiento.

7.5.6 Estudio de caso

El *estudio de caso* es una modalidad investigativa que se utiliza ampliamente, con excelentes resultados desde inicios del siglo XXI, en las ciencias sociales, en especial en la psicología, la educación, la salud, la sociología, la antropología y, de manera más reciente, en los campos de la economía y la administración.

www.FreeLibros.me

El objetivo de los estudios de caso, mejor conocido como el método del caso, es estudiar en profundidad o en detalle una unidad de análisis específica, tomada de un universo poblacional.

Para el método del caso, la unidad de análisis, "el caso" objeto de estudio es comprendido como un sistema integrado que interactúa en un contexto específico con características propias. El caso o unidad de análisis puede ser una persona, una institución o empresa, un grupo, etcétera.

Como método o procedimiento metodológico de investigación, el estudio de caso se desarrolla mediante un proceso cíclico y progresivo, el cual parte de la definición de un(os) tema(s) relevante(s) que se quiere(n) investigar. Se estudian en profundidad estos temas en la unidad de análisis, se recolectan los datos, se analizan, interpretan y validan; luego, se redacta el caso.

Los estudios de caso, como método de investigación, involucran aspectos descriptivos y explicativos de los temas objeto de estudio, pero además utilizan información tanto cualitativa como cuantitativa. También, como afirma Cerda (1998), aunque estos estudios ponen énfasis en el trabajo de campo, es imprescindible contar con un marco de referencia teórico, relacionado con los temas relevantes que los guían, para analizar e interpretar la información recolectada.

Para los expertos en investigación, los estudios de caso enfrentan la dificultad de no poder generalizar los resultados obtenidos de la unidad o el caso objeto de estudio, aunque hay ejemplos de teorías de amplios reconocimiento y validez generados a partir del estudio de muy pocos casos.

Las principales fuentes para la obtención de la información, en el estudio de caso, son las personas directamente relacionadas con el caso o la unidad de análisis y documentos de toda índole válidos que contengan información sobre el caso. Las técnicas más utilizadas y adecuadas para el estudio de caso son la observación estructurada, las entrevistas, los cuestionarios, los diarios, las autobiografías, los documentos personales, la correspondencia, etcétera.

Es importante no confundir el estudio de caso como "método de investigación" con el estudio de caso como "método o estrategia pedagógica de enseñanza-aprendizaje". Recuérdese que el estudio de caso, como método de investigación, es un procedimiento metodológico para estudiar en profundidad y en detalle una unidad de análisis dentro de un universo poblacional a partir de un(os) tema(s) de interés del investigador, siendo estos temas relevantes los que en todo momento guían el estudio.

El estudio de caso, como estrategia pedagógica del proceso enseñanza-aprendizaje, utilizado desde la antigüedad, consiste en presentarle al aprendiz o estudiante información sobre una situación real de un caso (empresa, organización, grupo o persona), para su análisis, reflexión y propuesta de alternativas de solución a la situación presentada y aplicable al caso al que hace relación la información. El uso de esa técnica o estrategia pedagógica favorece el desarrollo de habilidades como capacidad de análisis, síntesis y evaluación de la

www.FreeLibros.me

información; a la vez, posibilita el desarrollo del pensamiento crítico y el trabajo en equipo, pero de igual manera afirma rasgos personales como actitudes y valores éticos, si el caso es bien conducido.

EJEMPLO 7.17 **Algunos temas de estudio de casos**

En psicología y educación

- Estudiar la relación entre factores medioambientales, heredados, tomando como unidad de análisis a una familia.
- Analizar los factores desencadenantes de conductas delincuenciales en adolescentes de un barrio de una ciudad, tomando como objeto de estudio un barrio determinado.
- Analizar la relación entre el grado de afecto recibido por los hijos de sus padres y el rendimiento académico, mediante el estudio de una familia con hijos en edad escolar.
- Estudiar los efectos emocionales que hay en las familias en las cuales uno de sus miembros tienen una enfermedad terminal.

En economía y administración

- Estudiar la innovación empresarial, tomando como referencia una o varias empresas, cada una como un caso.
- Estudiar el estilo de dirección en una empresa.
- Estudiar el grado de competitividad de una organización o un sector como caso.

7.5.7 Investigaciones experimentales

En los tipos de investigación mencionados (histórica, documental, descriptiva, correlacional, estudios de caso, explicativa), así como en los estudios seccionales y longitudinales que se presentarán más adelante, el investigador es un observador meticuloso y sistemático que obtiene información del objeto de estudio mediante encuestas, entrevistas, documentos, análisis de vestigios, etcétera, que luego procesa e interpreta, con lo cual redacta un informe pero no ejerce ninguna acción directa que implique actuar sobre el objeto de estudio para conocer la información producida como consecuencia del acto ejercido por él.

La investigación experimental se caracteriza porque en ella el investigador actúa conscientemente sobre el objeto de estudio, en tanto que los objetivos de estos estudios son precisamente conocer los efectos de los actos producidos por el propio investigador como mecanismo o técnica para probar sus hipótesis.

De acuerdo con los expertos y seguidores del paradigma positivista, la experimentación es el verdadero método o el método por excelencia de la investigación científica; por tanto, la investigación experimental es la verdadera investigación, y el conocimiento generado por ésta es el conocimiento realmente válido y científico.

www.FreeLibros.me

Para muchos expertos en investigación, la investigación experimental es más propia de las ciencias naturales que de las ciencias sociales, debido a que por razones de carácter axiológico (ético) se cuestionan las implicaciones que ésta pueda tener en el ser humano en el acto de manipular, provocar y ejercer acciones por los investigadores, quienes lograrían reducir al ser humano al nivel de un objeto manipulable.

Muchos de los experimentos en psicología, educación y medicina se cuestionan porque es usual hallar que la investigación, que debería tener como fin servir al ser humano, convierte al ser humano en instrumento y objeto al servicio de la misma ciencia. Quienes realicen investigación experimental deben estar siempre conscientes de tener al ser humano como primer beneficiario de los resultados de la ciencia, y no como un medio más al servicio de ésta.

Entonces, ¿qué es la investigación experimental? De acuerdo con Hernández Pina et al. (2001) un experimento es un proceso planificado de investigar en el que al menos una variable (llamada experimental o independiente: VI) es manipulada u operada intencionalmente por el investigador para conocer qué efectos produce ésta en al menos otra variable llamada dependiente (VD). Para la mencionada autora, la variable independiente se conoce también como variable experimental o de tratamiento; la variable dependiente, conocida también como resultados o efecto, se refiere a los efectos observados en el estudio. Además de las variables independientes (VI) y las variables dependientes (VD) hay otras variables que actúan durante el experimento, las cuales se denominan variables intervinientes y extrañas, que en un experimento puro, verdadero, es necesario controlar para reducir su efecto, lo cual se hace creando unas condiciones especiales de laboratorio y de escogencia de sujetos de la investigación.

Toda investigación experimental parte de hipótesis, por lo que el objetivo es probar esa hipótesis.

7.5.8 Otros tipos de investigación

Según Briones (1985), en relación con el tiempo o número de veces en que se obtiene información del objeto de estudio, existen dos tipos de investigación: la seccional y la longitudinal.

Investigaciones seccionales o transversales

Son aquellas en las cuales se obtiene información del objeto de estudio (población o muestra) una única vez en un momento dado. Según Briones (1985), estos estudios son especies de "fotografías instantáneas" del fenómeno objeto de estudio. Esta investigación puede ser descriptiva o explicativa.

www.FreeLibros.me

EJEMPLO 7.18 **Algunas investigaciones seccionales o transversales**

En el campo de las organizaciones

- Las encuestas realizadas a una población o muestra de empresarios en una determinada fecha.
- Las encuestas realizadas a un grupo de consumidores en un determinado momento.

En psicología y educación

- Un estudio en donde se quieren conocer las actitudes y competencias creativas en las distintas edades del desarrollo humano, para lo cual se toman personas (hombres y mujeres) de distintas edades, y en el mismo período se hace la prueba o evaluación de tales variables.
- Un estudio que busca conocer las características de cada uno de los ciclos de la pareja, desde el noviazgo hasta la muerte o la separación de los miembros, toma parejas en distintos momentos e identifica los rasgos propios de cada ciclo de vida de la pareja.

Investigaciones longitudinales

A diferencia de la investigación seccional que obtiene datos de un objeto en una sola oportunidad, en la *investigación longitudinal* se obtienen datos de la misma población en distintos momentos durante un período determinado, con la finalidad de examinar sus variaciones en el tiempo.

EJEMPLO 7.19 **Algunas investigaciones longitudinales**

En el campo de las organizaciones

- Un estudio orientado a evaluar la productividad de los trabajadores de una empresa donde se analiza la producción durante dos años, recogiendo información tres veces por año.

En psicología y educación

- Un estudio que pretende conocer la disposición al cambio de las personas según la edad, toma a una o varias personas de la misma edad y evalúa la disposición del cambio de las mismas a medida que pasan por distintas edades.
- Un estudio que busca conocer el grado de autonomía en las decisiones de los estudiantes según años de estudio. Puede hacerse con personas desde los últimos años del bachillerato y hacer seguimiento de las mismas para evaluar su capacidad de autonomía a medida que avanzan en sus años de estudio.

Finalmente, y para ilustrar mejor a los lectores, actualmente los debates filosófico-científicos han adoptado nominaciones de acuerdo con los paradigmas y las epistemologías

www.FreeLibros.me

de tipos de investigación como: investigación ideográfica, investigación evaluativa, investigación etnometodológica, investigación naturalista, investigación histórico-hermenéutica, investigación fenomenológica, investigación crítica-interpretativa, investigación básica, investigación aplicada, investigación para el desarrollo, etcétera.

La elección del tipo de investigación por parte del investigador depende de tres factores: el objeto, el problema de investigación, los objetivos y el tipo de hipótesis, si es que se plantea por parte del investigador, por ello se deben tener muy claros tales aspectos o factores del proceso investigativo.

La Tabla 7.1 muestra una síntesis de los principales tipos de investigación, según las características más relevantes de cada uno de éstos.

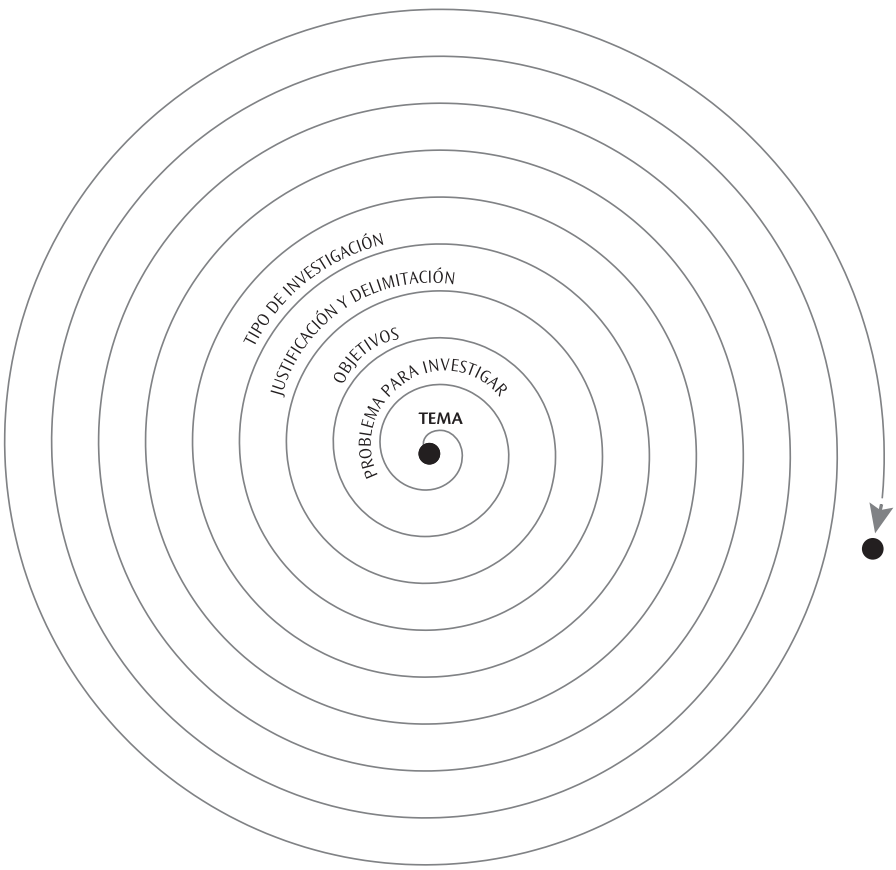
TABLA 7.1 **Tipos de investigación**

Investigación	Características
Histórica	Analiza eventos del pasado y busca relacionarlos con otros del presente.
Documental	Analiza información escrita sobre el tema objeto de estudio.
Descriptiva	Reseña rasgos, cualidades o atributos de la población objeto de estudio.
Correlacional	Mide el grado de relación entre variables de la población estudiada.
Explicativa	Da razones del porqué de los fenómenos.
Estudio de caso	Analiza una unidad específica de un universo poblacional.
Seccional	Recoge información del objeto de estudio en oportunidad única.
Longitudinal	Compara datos obtenidos en diferentes oportunidades o momentos de una misma población, con el propósito de evaluar cambios.
Experimental	Analiza el efecto producido por la acción o manipulación de una o más variables independientes sobre una o varias dependientes.

La Figura 7.8 ilustra la relación del tipo de investigación con los componentes antes mencionados, según la espiral del proceso de investigación propuesto en el inicio de esta parte II del libro.

www.FreeLibros.me

FIGURA 7.8 **El proceso de investigación. Definición del tipo de investigación**



RESUMEN

En investigación existen diferentes tipos de estudio y es necesario conocer sus características para saber cuál de ellos se ajusta mejor a la investigación que va a realizarse. La elección del tipo de investigación depende, en alto grado, del objetivo del estudio del problema de investigación y de las hipótesis que se formulen en el trabajo, así como de la concepción epistemológica y filosófica de la persona o del equipo investigador.

Los principales tipos de estudio presentados en este capítulo son:

www.FreeLibros.me

La *investigación histórica*, que se orienta a estudiar los sucesos del pasado. Analiza la relación de esos sucesos con otros eventos de la época y con sucesos presentes. En síntesis, busca entender el pasado y su relación con el presente y el futuro. Esta investigación tiene como sus principales fuentes y técnicas de obtención de la información, la revisión documental, los vestigios y objetos reales y, en algunos casos, personas que tuvieron relación directa con los hechos estudiados o quienes, aunque no tuvieron relación directa con esos hechos, cuentan con información válida y confiable sobre los mismos.

La *investigación documental*, que consiste en un análisis de la información escrita sobre un determinado tema, con el propósito de establecer relaciones, diferencias, etapas, posturas o estado actual del conocimiento respecto al tema objeto de estudio. Las principales fuentes de información en este tipo de investigación son: documentos escritos (libros, periódicos, revistas, actas notariales, tratados, conferencias escritas, etcétera), documentos fílmicos (películas, diapositivas, etcétera) y documentos grabados (discos, cintas, casetes, disquetes, etcétera).

La *investigación descriptiva*, aquella que reseña las características o los rasgos de la situación o del fenómeno objeto de estudio. Es uno de los tipos o procedimientos investigativos más populares y utilizados por los principiantes en la actividad investigativa. La realización de este tipo de investigación se soporta principalmente en técnicas como la encuesta, la entrevista, la observación y la revisión documental.

La *investigación correlacional*, que tiene como propósito mostrar o examinar la relación entre variables o resultados de variables. Examina relaciones entre variables o sus resultados, pero en ningún momento explica causas o consecuencias. Su principal soporte es el uso de herramientas estadísticas.

La *investigación explicativa o causal*, para muchos expertos es el ideal y el culmen de la investigación no experimental, tiene como fundamento la prueba de hipótesis y busca que las conclusiones lleven a la formulación o al contraste de leyes o principios científicos. En la investigación explicativa se analizan causas y efectos de la relación entre variables.

El *estudio de caso*, también conocido como el método del caso, consiste en estudiar en profundidad o en detalle una unidad de análisis específica, tomada de un universo poblacional. El caso o unidad de análisis puede ser una persona, una institución o empresa, un grupo, etcétera. Las principales fuentes para la obtención de la información, en el estudio de caso, son las personas directamente relacionadas con el caso o la unidad de análisis y documentos de toda índole válidos que contengan información sobre el caso. Las técnicas más utilizadas y adecuadas para el estudio de caso son la observación estructurada, las entrevistas, los cuestionarios, los diarios, las autobiografías, los documentos personales, la correspondencia, etcétera.

www.FreeLibros.me

La *investigación experimental*, que se caracteriza porque en ella el investigador actúa conscientemente sobre el objeto de estudio, en tanto que los objetivos de estos estudios son precisamente conocer los efectos de los actos producidos por el propio investigador como mecanismo o técnica para probar sus hipótesis. Es el método por excelencia de la investigación científica; por tanto, para muchos expertos en el tema, la investigación experimental es la verdadera investigación.

La *investigación seccional o transversal*, aquella en la cual se obtiene información del objeto de estudio (población o muestra) una única vez en un momento dado.

La *investigación longitudinal* que, a diferencia de la investigación seccional que obtiene datos de un objeto en una sola oportunidad, obtiene datos de la misma población en distintos momentos durante un período determinado, con la finalidad de examinar sus variaciones en el tiempo.

Finalmente, vale recordar que la elección del tipo de investigación por el investigador depende de tres factores básicos: el objeto, el problema de investigación, los objetivos y el tipo de hipótesis, si es que se plantea por parte del investigador.

EJERCICIOS DE REPASO Y ANÁLISIS

1. Enuncie y comente los principales tipos de investigación presentados en este apartado.
2. Explique la relación entre el problema de investigación, los objetivos y tipo de investigación en el desarrollo de una investigación científica.
3. Consulte en bibliotecas y medios especializados en material científico y obtenga información relacionada con ejemplos de investigaciones recientes en el campo de su profesión, según los diferentes tipos de investigación. Elabore un resumen de cada una de las investigaciones consultadas.
4. Proponga ejemplos de temas de investigación histórica, documental, descriptiva, correlacional, explicativa, longitudinal, transversal y de estudio de caso en el campo de su profesión, que usted considere que merecen desarrollarse como trabajos de grado.
5. Averigüe qué otros tipos de investigación distintos de los presentados en este apartado existen en el campo de la investigación científica y comente qué pertinencia tienen para el campo de su profesión.
6. En trabajo en equipo con sus compañeros de clases, compartan la información de los puntos 1, 2, 3, 4 y 5. Elaboren un informe de la actividad del intercambio de información.

www.FreeLibros.me

7.6 Marco de referencia de la investigación

Como la ciencia es una búsqueda permanente del conocimiento válido, cada nueva investigación debe fundamentarse en el conocimiento existente y de igual manera asumir una posición frente al mismo. Por este motivo, toda investigación debe realizarse dentro de un *marco de referencia* o conocimiento previo, es decir, es necesario ubicar la investigación que va a realizarse dentro de una teoría, un enfoque o una escuela. También se debe explicar la concepción de persona que enmarcará la investigación y, finalmente, se deben precisar los conceptos relevantes del estudio.

En síntesis, el marco de referencia es el marco general de la fundamentación teórica y antropológica en la cual se desarrolla el estudio. Este marco comprende:

- El marco filosófico-antropológico (opcional).
- El marco teórico.
- Otros marcos (legal, histórico, etcétera).

En algunos casos debe considerarse, además, el marco legal.

7.6.1 Marco filosófico-antropológico

Consiste en exponer la concepción filosófico-antropológica o concepto de ser humano que tienen el autor o los autores de la investigación. Ese concepto se reflejará en el desarrollo del estudio, enfatizando algunos aspectos y relegando otros.

Al referirse a la importancia de las suposiciones, Drucker (1999) afirma: "Las suposiciones básicas que una disciplina mantiene acerca de la realidad determinan el punto de concentración de la disciplina, decidiendo qué va a recibir atención y qué va a desconocer o pasar por alto" (p. 3). En el campo específico de las ciencias económicas y administrativas, las suposiciones o concepciones respecto al ser humano son fundamentales debido a que éstas han determinado la actitud que los gobernantes y directivos empresariales han asumido en torno al ser humano, ya sea para mejorar su calidad de vida y promover su desarrollo integral, o para contribuir a la deshumanización, cosificación o marginación de la persona.

En ese sentido, conscientes de que la investigación debe, ante todo, orientarse en pro del ser humano, es importante que en un proceso de investigación se explicite la concepción del ser humano que fundamenta el estudio respectivo, generándose compromiso del investigador de asumir lo humano como lo fundamental.

¿Cómo elaborar el marco filosófico-antropológico de una investigación que se va a realizar?

Para la elaboración del marco filosófico-antropológico de una investigación es necesario revisar bibliografía específica relacionada con la concepción del ser humano o persona, que tenga directa relación con el tema y el problema de investigación, pero donde se resalte lo

www.FreeLibros.me

humano en sentido positivo como lo fundamental. Para ello es imprescindible consultar a expertos en temas antropológico-filosóficos o en humanidades, quienes guiarán en cuanto a referencias bibliográficas y orientarán en la construcción de un marco antropológico sobre el tema objeto de la investigación.

En la elaboración de este marco, se revisa la información encontrada sobre aspectos relacionados con el concepto antropológico y filosófico del ser humano, luego se redacta una argumentación sobre la concepción de persona que fundamentará el estudio. En el ejemplo 7.20 se muestra cómo se presenta un marco filosófico-antropológico.

7.6.2 Marco teórico

De acuerdo con Hugo Cerda (1998), "es imposible concebir una investigación científica sin la presencia de un marco teórico, porque a éste le corresponde la función de orientar y crear las bases teóricas de la investigación" (p. 170). Según esta afirmación, ¿qué es, entonces, el marco teórico en una investigación científica?

Definición

Si bien no existe consenso entre los distintos investigadores respecto a qué se entiende por marco teórico, según Briones, citado por Cerda (1998):

El marco teórico a niveles más específicos y concretos comprende la ubicación del problema en una determinada situación histórico-social, sus relaciones con otros fenómenos, las relaciones de los resultados por alcanzar con otros ya logrados, como también definiciones de nuevos conceptos, redefiniciones de otros, clasificaciones, tipologías por usar, etcétera. (p. 171).

El marco teórico se entenderá aquí como la fundamentación teórica dentro de la cual se enmarcará la investigación que va a realizarse. Es decir, es una presentación de las principales escuelas, enfoques o teorías existentes sobre el tema objeto de estudio, en que se muestre el nivel del conocimiento en ese campo, los principales debates, resultados, instrumentos utilizados, y demás aspectos pertinentes y relevantes sobre el tema de interés.

El marco teórico no es un resumen de las teorías que se han escrito sobre el tema objeto de la investigación; más bien es una revisión de lo que está investigándose o se ha investigado en el tema objeto de estudio y los planteamientos que sobre el mismo tienen los estudiosos de éste. Esta fundamentación soportará el desarrollo del estudio y la discusión de los resultados.

Como la cantidad de información en algunos temas abunda, es importante que la persona interesada en desarrollar cualquier investigación se asesore de expertos en ese campo, y se concentre en la información pertinente y relevante del tema. Hay que leer de preferencia aquellos documentos y libros especializados que muestran los resultados de las últimas investigaciones realizadas sobre el tema, para así evitar desperdiciar recursos.

Con los nuevos sistemas de comunicación vía Internet, es más fácil obtener y aprovechar información sobre cualquier tema en cualquier parte del mundo.

www.FreeLibros.me

Funciones del marco teórico

De acuerdo con Tamayo y Tamayo (2002), el marco teórico de una investigación tiene, entre otras, las siguientes funciones:

- Sirve de base para la descripción o el enunciado del problema. Integra la teoría con la investigación y sus relaciones mutuas.
- Sirve para cuestionar, así como para conducir a una reformulación e incluso a un cambio del problema de investigación.
- Ayuda a precisar y a organizar los elementos contenidos en la descripción del problema.
- Delimita el área de la investigación.
- Sugiere guías de investigación, es decir, con el marco teórico pueden verse nuevas alternativas de enfoque para tratar el problema.
- Compendia conocimientos en el área que se va a investigar.
- Expresa proposiciones teóricas generales, postulados y marcos de referencia que sirven como base para formular hipótesis, manejar variables y esbozar teoría sobre técnicas y procedimientos por seguir en el estudio que va a realizarse.

El marco teórico, fundamento de la discusión de resultados de la investigación, se constituye en un aspecto crucial para la redacción de las conclusiones del informe final del estudio.

Elaboración del marco teórico

Si bien no existe una guía específica para la elaboración de un marco teórico para una investigación, se recomienda comenzar por revisar libros, revistas y demás documentos especializados que aborden el tema que se va a investigar; para ello, se sugiere recurrir a la asesoría de expertos en ese tema, quienes orientarán la consecución de la bibliografía más adecuada.

Con base en la bibliografía revisada, se debe llegar a un conocimiento amplio y detallado del estado del tema, mostrando las investigaciones que se han adelantado y están adelantándose, los objetivos y las hipótesis que han guiado dichas investigaciones y las conclusiones a que se ha llegado, la relevancia que tiene el tema en la actualidad, las posturas y los enfoques que abordan el tema, los instrumentos y las técnicas de recolección de información que se han utilizado en los estudios, etcétera.

Con los resultados de la revisión bibliográfica se construye el marco teórico, integrando las ideas y los resultados de los distintos documentos revisados, siguiendo algún criterio metodológico según temas o aspectos tratados en esa bibliografía, que tengan relación con el objeto de estudio y en particular con el problema de investigación, los objetivos y las hipótesis por probar, si las hubiere.

Por último, se enmarca la investigación por realizar dentro de este marco, ya sea desde la adopción de un determinado enfoque de los identificados en la revisión bibliográfica, un modelo o una técnica de recolección de la información, que guiará el desarrollo del estudio.

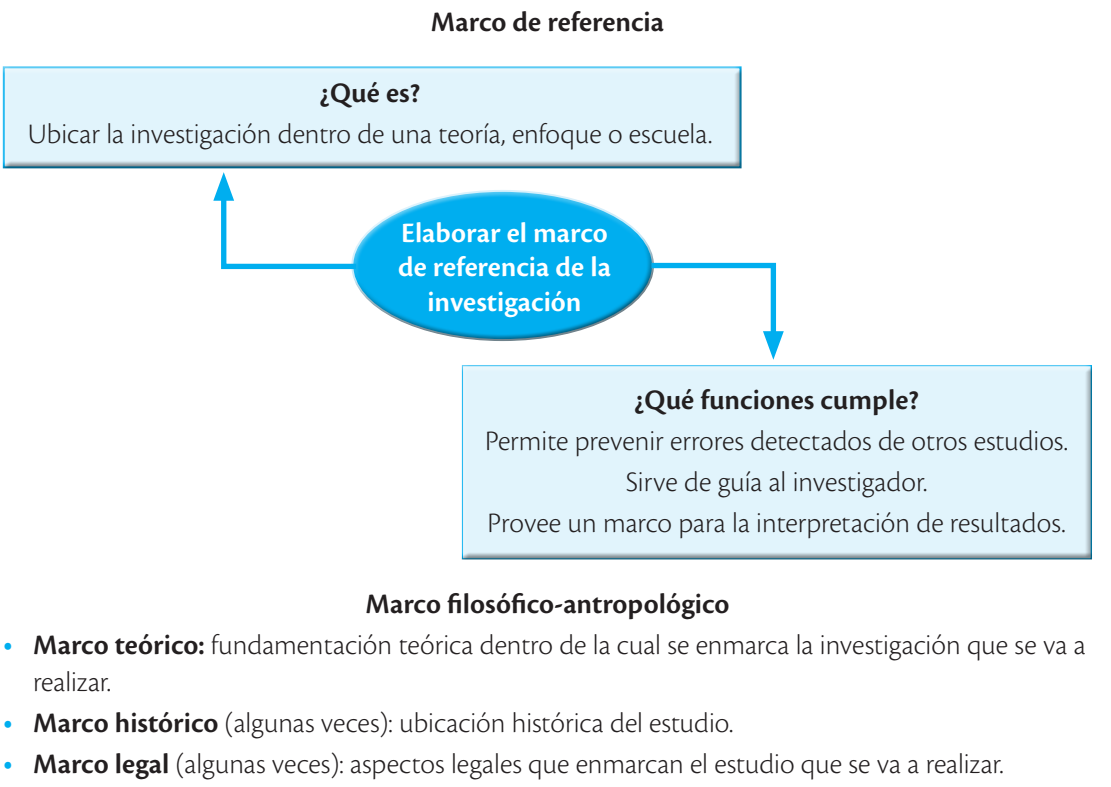
www.FreeLibros.me

Es importante aclarar que para la propuesta o el anteproyecto de la investigación que se va a realizar no es necesario contar con un marco teórico detallado y demasiado riguroso, pero sí se requiere tener un marco teórico básico, donde se presenten las ideas y los aspectos relevantes que en el momento hay sobre el tema de la investigación. Para el informe final de la investigación sí es indispensable contar con un marco teórico sólido, donde se muestre en detalle el estado del conocimiento sobre el tema en cuestión y la relevancia del mismo en el área del conocimiento.

Vale recordar que, en la elaboración del marco teórico, es necesario elaborar las citas bibliográficas y las notas de pie de página (*Normas de Chicago*, *Normas APA* o las normas técnicas de presentación de documentos de investigación específicas de cada país donde se presentará el respectivo documento).

La Figura 7.9 sintetiza el concepto y las funciones del marco de referencia en una investigación junto a los diferentes submarcos (filosófico-antropológico, teórico, legal, etcétera).

FIGURA 7.9 **Concepto y funciones del marco de referencia**



www.FreeLibros.me

EJEMPLO 7.20 **Marcos de referencia en una investigación**

En el campo organizacional

Título de la investigación: *Creatividad de innovación en las organizaciones*

Marco de referencia

Con el propósito de ambientar al lector, a continuación se presentan los fundamentos antropológicos y teóricos que enmarcan este estudio y que sirven de base para el análisis de sus resultados.

• Marco filosófico-antropológico

El presente trabajo está enmarcado en una concepción de ser humano como una entidad compleja y paradójica y, sobre todo, como un ser pluridimensional (Polo, 1997), ya que es un ser en sí mismo, es decir, un ser en relación con los otros seres humanos, un ser corpóreo, espiritual, emocional, intelectual, cultural e histórico, libre, trascendente, y un ser en proyecto, que se construye y transforma el mundo en que vive. Un ser capaz de perfeccionarse es un ser creativo y optimista para formarse una vida digna y humana. Es un ser afectivo, que sufre y goza. Un ser con limitaciones, pero dispuesto a superarlas. Es un ser capaz de mejorar cada día con la ayuda de los demás.

• Marco teórico

Siguiendo el ejemplo de la creatividad en las organizaciones, el siguiente es un ejemplo resumido de un marco teórico:

La investigación en los distintos campos del conocimiento humano cada día concede más importancia y atención a cuanto implica la actividad creadora, por la convicción de que el desarrollo de la persona humana, de las organizaciones y de la sociedad en general se da en proporción directa con su poder creador o cocreador (López Quintás, 1998, p. 89) .

Instituciones como las Naciones Unidas y científicos reconocidos en distintos campos del conocimiento no dudan en afirmar que sólo mediante la creatividad se garantizarán el desarrollo integral y el progreso de la humanidad en los próximos años. Por ello ésta debe ser una preocupación de líderes nacionales y directivos organizacionales en los países en vía de desarrollo, como lo afirma Phillips (1994). Sin embargo, para Sternberg y Lubart (1997), la creatividad y la innovación son menospreciadas por la sociedad, en general, y por las instituciones que existen dentro de la sociedad. Estos investigadores afirman que los ejecutivos del mundo de los negocios hablan de la necesidad de la creatividad y de la innovación; pero sólo lo hacen en el sentido demagógico, pues en la realidad no la promueven.

Otros investigadores como Finke (1998) aseguran que los resultados de la dinámica de las organizaciones impresionan más por lo lento de sus cambios que por la rapidez y novedad de ellos. Finke (1998) afirma que "las culturas organizacionales y los modos de hacer las cosas parecen tener una vida que se prolonga más allá de las personas particulares que habitan la

www.FreeLibros.me

organización" (p. 321). Por otro lado, Gardner (1997) muestra que la creatividad es algo difícil de encontrar en el mundo de los negocios, tal vez porque en gran medida los ejecutivos reconocen su necesidad, aunque, por distintos motivos, le temen.

Por ello, Sternberg y Lubart (1997) afirman que a pesar de que muchas personas dicen valorar las ideas originales, existen pruebas contundentes de que no están identificadas con aquello que supuestamente valoran. Al respecto, uno de los hallazgos más consecuentes en psicología es el muy conocido "efecto de mera exposición", planteado por Zajona en 1978 y que consiste en mostrar que a las personas, en su mayoría, les gusta lo que les es familiar. Con esto puede afirmarse que aunque mediante palabras podría valorarse la creatividad, porque trae consigo el progreso, muy a menudo incomoda a la gente y, en consecuencia, quienes no la practican reaccionarían negativamente ante la persona o la obra creativa.

Puesto que en las organizaciones los directivos parecen poco creativos, Kuczmarski (1997), afirma que las personas creativas a menudo se consideran raras e incluso se les margina.

Dados los rápidos y complejos cambios que vive el mundo empresarial, resulta oportuno conocer cómo se percibe y se actúa en las empresas nacionales en cuanto a la creatividad. En el país se habla insistentemente de la necesidad de dinamizar las organizaciones para acomodarse a las exigencias del nuevo ambiente de los negocios; por ello, conviene explorar la manera como se asume este reto, y se aprovecha la capacidad innovadora y creativa del recurso humano en las diferentes organizaciones.

EJEMPLO 7.21 **Marco teórico en una investigación**

En temas de educación y psicología

Título de la investigación: *"Compromiso pedagógico de la literatura infantil latinoamericana producida en la década de 1990 en la formación integral del niño y la niña".*

Marco de referencia

El siguiente es el marco general que guía este estudio: se comienza con un marco antropológico-filosófico cuyo propósito es mostrar la concepción del "ser niño(a)", que fundamenta el respectivo estudio; luego, se presenta el marco teórico que soporta el tema específico objeto del estudio.

- ***Marco filosófico-antropológico***

Conscientes de las implicaciones que toda investigación llega a tener sobre el sujeto de estudio, ya sea por la concepción epistemológica, el método de investigación utilizado o la visión antropológica que los investigadores tengan del mundo, la vida y del propio ser humano, pero en particular del sujeto específico de la investigación, a continuación se presenta, de forma breve, el marco filosófico-antropológico o concepción de niña(o) que guía el estudio.

www.FreeLibros.me

A lo largo de las diferentes épocas, las personas han sostenido diversas concepciones acerca de lo que son los niños y las niñas, así como de la manera en la que deben ser educados. Según Arias, citado por Papalia (2000), sólo a partir del siglo XVII los niños han sido vistos como seres humanos diferentes a los adultos, ya que, de acuerdo con Arias, en épocas anteriores, a los infantes se les consideraba adultos pequeños, débiles y menos inteligentes que los adultos.

Sin embargo, investigaciones recientes realizadas por Elkin (1987), citado por Hurlock (2003), muestran que los antiguos griegos y romanos daban un reconocimiento especial a la naturaleza de los niños y las niñas. Igualmente, Pollack (1989), citado por el mismo Hurlock (2003), afirma que los "niños siempre han sido vistos y tratados de manera diferente a la de los adultos" (p. 173).

No obstante las anteriores afirmaciones, hay acuerdo, a nivel educativo, de que una de las primeras personas que comienza a concebir, estudiar y educar a los niños y niñas como personas con rasgos propios diferentes a los de los adultos fue el famoso reformador eslavo de la educación del siglo XVII: John Amos Comenius. Posteriormente, y en la misma dirección de Comenius, estudiosos de los niños y niñas tales como Lock, Rousseau, Pestalozzi, Herbart y Frocbel dieron origen a tratados filosóficos y antropológicos para la comprensión y el estudio de los niños y las niñas.

Sólo a partir del siglo XIX, gracias a la proliferación de estudios sistemáticos sobre la niñez como una etapa de vida del ser humano con características propias, diferentes a las del adulto, se comienza a tener plena conciencia del significado del ser niño(a) o adulto, y la filosofía y la antropología se encargan de mostrar que esos significados cambian en relación con los tiempos, lugares, papeles y derechos asignados a uno y a otro; es así como desde la última década la categoría de infancia ha venido convirtiéndose en tema crucial para la educación y la sociedad.

Con la finalidad de motivar la atención mundial en beneficio de la infancia, 1979 se declara el "Año Internacional del Niño", en tanto que en 1989, en la Asamblea General de las Naciones Unidas, se proclama la "Convención Internacional sobre los Derechos del Niño y la Niña" y se inicia una nueva conceptuación referente a la protección y al desarrollo integral de la niñez, fundamentada en una concepción de niños y niñas como personas en proceso de formación, que necesitan del apoyo de los adultos y cuya formación se debe orientar hacia la autonomía y la libertad.

En virtud de lo mencionado, ha venido trabajándose en pro de mejorar la situación de la niñez latinoamericana, en razón de la transformación histórica de múltiples procesos sociales, económicos, políticos y culturales en los contextos locales y globales, así como del conocimiento sobre la naturaleza del niño y la niña, sobre sus procesos de desarrollo en sus diferentes dimensiones como ser integral que configuran su potencial y su identidad, lo mismo que sobre los procesos de intervención que puedan contribuir a acelerar y cualificar su desarrollo integral.

Ahora, conocedores de que el desarrollo del ser humano, y por ende el de la niñez, es resultado de la interacción entre la carga genética y el entorno físico y social que lo rodea desde su nacimiento, el desarrollo del niño y la niña es una construcción que exige la participación tanto de ellos(as) como de aquellos que interactúan con ellos, según lo afirma Coll (1990), citado por Bassedas et al. (2000). Así que una relación positiva y constructiva entre los adultos y los niños

www.FreeLibros.me

y las niñas es imprescindible para el aprendizaje que potencie el desarrollo integral y armónico de estos últimos.

Las personalidades del niño y la niña se estructuran a partir de la relación entre éstos y el entorno, pero principalmente en la relación con otras personas, por lo cual es lógico, como afirma Bassedas et al. (2000), atribuir gran importancia a las relaciones y las interacciones entre las personas como factor facilitador de la construcción progresiva de la identidad personal y del desarrollo de las diferentes potencialidades de toda persona, especialmente en la etapa de la niñez, que es la etapa cuando las personas interiorizan la imagen de sí mismos y toman los patrones sociales para valoración del concepto que construyen de los demás.

De acuerdo con lo comentado, y dado que la literatura es un medio de contacto usual de los niños o las niñas, la literatura debe ser un medio a través del cual los niños y las niñas puedan forjarse una conciencia de sí mismos y del mundo que los rodea, de sus sentimientos, de sus actitudes éticas y de sus juicios de valor. Al respecto, Rosenblatt (1998) considera que la literatura ejerce una gran influencia en el desarrollo personal, social e intelectual del niño y la niña; además, dado el poder que tienen los textos literarios sobre el lector, al ofrecer experiencia directa de los problemas con los que se enfrenta el ser humano, es determinante en la formación de su personalidad.

En síntesis, para una concepción de niño(a) como ser pluridimensional, autónomo y libre, con identidad propia y diferente a la del adulto, se requiere una literatura que contribuya al desarrollo integral de sus potencialidades y a la formación de un ciudadano que coadyuve a construir una mejor sociedad fundamentada en valores éticos, estéticos y humanos, ya que la niñez es considerada hoy como el eje del desarrollo social, cultural y político de la sociedad.

• *Marco teórico*

Con base en el marco antropológico definido, a continuación se hace una presentación del marco teórico que orienta el tema específico del estudio, el cual servirá de base para el trabajo de campo y la discusión de resultados obtenidos una vez realizado el mismo.

De acuerdo con Nobile (1992), aunque actualmente no hay consenso respecto a qué se entiende por literatura infantil, sí hay expresiones utilizadas con mayor o menor convicción y legitimidad para designar ese especial ámbito literario dirigido a los niños y las niñas.

Para Goreti (1968), citado por Nobile, la literatura infantil es aquella que se ha pensado expresamente y que va dirigida para la infancia. Sin embargo, estudiosos de la literatura como Bernardines y Bertordine (1998) afirman que la literatura infantil no es sólo la que los escritores escriben, sino también la que los niños y las niñas aceptan y hacen suya al leerla.

En efecto, la crítica actual, afirma Nobile (1992), tiende a englobar, en el concepto de literatura infantil, todo cuanto se ha escrito para la edad evolutiva y las obras de las que los niños y las niñas han ido apropiándose, sin excluir los mensajes narrativos ofrecidos por los nuevos instrumentos de comunicación. En este sentido, agrega la autora, literatura infantil es todo lo que ha sido y es narrado expresamente para ser oído y leído por los infantes y jóvenes como

www.FreeLibros.me

interlocutores activos, con "exclusión" de lo que, utilizando formas de literatura, se ha propuesto y se propone con unas finalidades distintas de las del diálogo cultural, libre y creativo que no es literatura. Como afirma González Gil (1993), no es posible llamar literatura a un libro que se mueve automáticamente y llora, a un libro que sólo huele o toca música.

El ámbito literario así delimitado, afirma Nobile (1992), comprende las obras narrativas destinadas a las edades infantil y juvenil, en sus distintas fases y estudios, que son capaces de promover una experiencia vital en el lector y de dirigirlo a un crecimiento, a una maduración global en todos los componentes de su personalidad.

La literatura infantil es, pues, una literatura en la que los dotes poéticos del escritor se conjugan para generar una viva sensibilidad educativa, capaz de descubrir horizontes de vida y nuevos conocimientos, ante la inagotable sed de saber y de exploración característica de la infancia, así como de enriquecer la experiencia humana y lingüística, representando una ventana abierta al mundo.

Para Gallego (1995), literatura infantil es el conjunto de obras de arte literario de diferentes épocas y lugares, escritos o narrados por adultos o infantes, con la intención o sin ésta de dirigirse a los niños y las niñas, pero que son aceptados por los mismos al encontrar en ellas una serie de características literarias, lingüísticas y sociales, de donde se logra no sólo el deleite, sino también el enriquecimiento integral de su personalidad o, en palabras de Gutiérrez (1989), la literatura infantil explora las particularidades y percepciones de la realidad, así como la representación estética de los chicos y las chicas, basándose en sus intereses y posibilidades intelectuales, la cual se rige por un equilibrio de valores estéticos y lúdicos, con el predominio de la palabra como recurso expresivo.

De acuerdo con lo mencionado, los escritores de literatura infantil deben tener una concepción de niño y niña como seres integrales, con dimensiones corporales, espirituales, intelectuales, psicoafectivas, sociales y axiológicas; por tanto, las obras literarias infantiles deben responder al lector como un ser libre, espiritual, cognoscente, autónomo, socioafectivo, ético y estético.

Rosenblatt (1998) considera que la literatura es un medio a través del cual los niños y las niñas llegan a forjarse una conciencia de sí mismos, del mundo que los rodea, de sus sentimientos, sus actitudes éticas y sus juicios de valor. Para la autora, la literatura tiene una gran influencia en el desarrollo personal, social e intelectual del lector, así como el poder que ejercen los textos literarios sobre el lector son determinantes en la formación de su personalidad, porque la literatura envuelve a la persona en su totalidad y abarca toda la condición humana de cada lector.

Según López Quintás (1990), no puede olvidarse que el lenguaje es el vehículo viviente de la construcción de ámbitos interpersonales, que se manifiestan de múltiples modos, por eso, además, resulta el medio por el cual los seres humanos crean ámbitos de convivencia, amistad, amor o conflicto y destrucción.

Para López Quintás (1990), la literatura debe ser un excelente medio para crear sentido de vida y construir realidades valiosas, donde se compromete a toda la persona, es decir, su voluntad, su entendimiento, su aspecto psicoafectivo, su espiritualidad, su dimensión estética y ética, su dimensión social y su dimensión corporal.

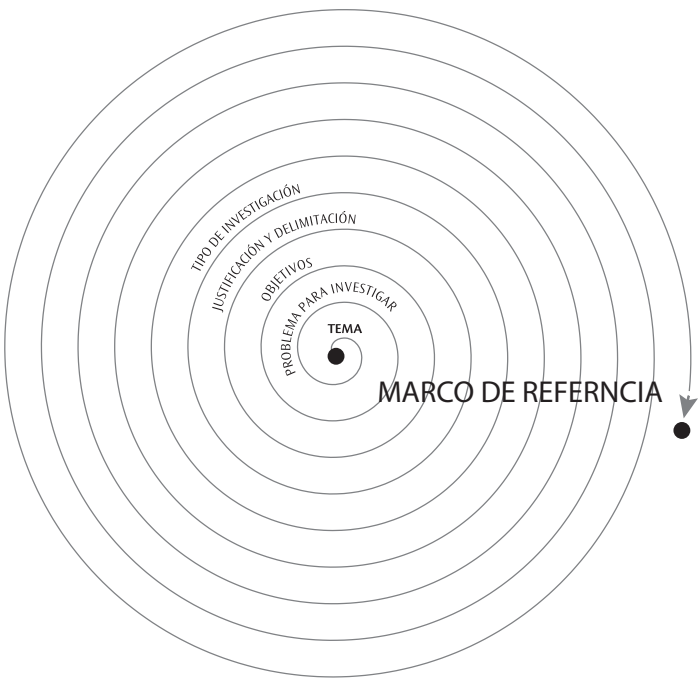
www.FreeLibros.me

¿La literatura infantil latinoamericana tiene un compromiso pedagógico orientado a la formación integral de sus lectores? Distintos autores e investigadores de la literatura infantil latinoamericana, como Hernán Rodríguez Castelo, narrador y crítico ecuatoriano, así como Silvio Castrillón, crítico de la literatura infantil, María Clemencia Venegas y Margarita Muñoz, de Colombia, lo mismo que Manuel Peña M., crítico e investigador literario cubano, concuerdan en afirmar que la literatura infantil latinoamericana ha tenido una acentuada intención pedagogizante, con tono moralizador y adoctrinador, pero no una orientación clara hacia la formación integral del niño o la niña.

Por este motivo, el presente estudio pretende conocer cuál ha sido la función pedagógica que promueve la literatura infantil latinoamericana producida durante la última década del siglo XX, según los expertos en el tema.

La Figura 7.10. muestra la relación entre el marco de referencia y los demás componentes del proceso de investigación, según el diagrama de espiral.

FIGURA 7.10 **El proceso de investigación. Elaboración del marco de referencia de la investigación**



www.FreeLibros.me

RESUMEN

La evidencia empírica de la propia ciencia indica que cada nueva investigación ha de realizarse dentro de un *marco de referencia* o conocimiento previo, es decir, es necesario ubicar la investigación que va a realizarse dentro de una teoría, un enfoque o una escuela. También, en algunas disciplinas, especialmente las relacionadas con la educación y el comportamiento humano, debe explicarse la concepción de la persona que enmarcará la investigación con el propósito de que el lector identifique el enfoque desde el cual se ha realizado la respectiva investigación. En síntesis, el marco de referencia es el marco general de la fundamentación teórica y antropológica en la cual se desarrolla el estudio.

El *marco de referencia* suele comprender varios marcos: el teórico, el filosófico-antropológico (opcional) y otros marcos (legal, histórico, etcétera).

El *marco filosófico-antropológico* consiste en exponer la concepción filosófico-antropológica o concepto de ser humano que tienen el autor o los autores de la investigación.

El *marco teórico* se entiende como la fundamentación teórica dentro de la cual se enmarcará la investigación que va a realizarse. Es decir, es una presentación de las principales escuelas, enfoques o teorías existentes sobre el tema objeto de estudio, en que se muestre el nivel del conocimiento en ese campo, los principales debates, resultados, instrumentos utilizados, y demás aspectos pertinentes y relevantes sobre el tema de interés.

El marco teórico, fundamento en la discusión de resultados de la investigación, se constituye en un aspecto fundamental para la redacción de las conclusiones del informe final del estudio. De otra parte, en el texto del marco teórico, es necesario mostrar las citas bibliográficas y las respectivas notas de pie de página y éstas deben presentarse según los manuales de estilo (Normas de Chicago, Normas APA o las normas técnicas de presentación de documentos de investigación específicas de cada país donde se presentará el respectivo documento).

Los marcos legal, histórico y demás dependen de las particularidades de cada investigación y su construcción será específica para cada proyecto.

www.FreeLibros.me

EJERCICIOS DE REPASO Y ANÁLISIS

1. ¿Qué es un marco de referencia en una investigación?
2. ¿Cuál es la importancia del marco antropológico en una investigación?
3. ¿Cuál es la relación entre el marco teórico y los objetivos de la investigación?
4. ¿Qué es el marco teórico y qué funciones cumple en una investigación?
5. Busque ejemplos publicados de marco teórico.
6. Seleccione un tema de investigación, establezca los objetivos y elabore el marco teórico, sin olvidar las referencias bibliográficas y las citas de pie de página. (Las normas de presentación de trabajos de grado se describen en el capítulo 8 de este libro). El marco teórico debe mostrar las principales teorías sobre el tema, las investigaciones más recientes y la relación con su investigación y el enfoque que desarrollará en la investigación, de acuerdo con la información consultada.
7. Con metodología de taller, lean los marcos teóricos elaborados por ustedes y analícenlos.

www.FreeLibros.me

7.7 Hipótesis de la investigación

7.7.1 Formulación de la hipótesis

Un aspecto importante en el proceso de investigación científica tiene que ver con las hipótesis, debido a que éstas son el medio por el cual se responde a la formulación del problema de investigación y se operacionalizan los objetivos.

Se formulan hipótesis cuando en la investigación se quiere probar una suposición y no sólo mostrar los rasgos característicos de una determinada situación. Es decir, se formulan hipótesis en las investigaciones que buscan probar el impacto que tienen algunas variables entre sí, o el efecto de un rasgo o una variable en relación con otro(a). Básicamente son estudios que muestran la relación causa/efecto.

Las investigaciones de tipo descriptivo no requieren formular hipótesis; es suficiente plantear algunas preguntas de investigación que, como ya se anotó, surgen del planteamiento del problema, de los objetivos y, por supuesto, del marco teórico que soporta el estudio.

En resumen, todo proyecto de investigación requiere preguntas de investigación, y sólo aquellos que buscan evaluar relación entre variables o explicar causas requieren la formulación de hipótesis. En el caso de la investigación experimental, siempre es necesario partir de hipótesis que serán las que guiarán el respectivo estudio.

7.7.2 Concepto de hipótesis

Según el Pequeño Larousse ilustrado, la palabra hipótesis "deriva del griego *hypotthesis*, que significa suposición de una cosa posible, de la que se saca una consecuencia" (García Pelayo, 1994, p. 544) .

Para Arias Galicia (1991), "una hipótesis es una suposición respecto de algunos elementos empíricos y otros conceptuales, y sus relaciones mutuas, que surge más allá de los hechos y las experiencias conocidas, con el propósito de llegar a una mayor comprensión de los mismos" (p. 66).

Para Muñoz Razo (1998), una hipótesis "es la explicación anticipada y provisional de alguna suposición que se trate de comprobar o desaprobar, a través de los antecedentes que se recopilan sobre el problema de investigación previamente planteado" (p. 94).

De las definiciones anteriores puede concluirse que una hipótesis es una suposición o solución anticipada al problema objeto de la investigación y, por tanto, la tarea del investigador debe orientarse a probar tal suposición o hipótesis. Ahora, es importante tener claro que al aceptar una hipótesis como cierta no se puede concluir respecto a la veracidad de los resultados obtenidos, sino que sólo se aporta evidencia en su favor.

www.FreeLibros.me

7.7.3 Función de las hipótesis

Las siguientes son algunas de las *funciones* que, según Arias Galicia (1991), cumplen las hipótesis en una investigación:

- Se precisan los problemas objeto de la investigación.
- Identifican o explicitan las variables objeto de análisis del estudio.
- Definen y unifican criterios, métodos, técnicas y procedimientos utilizados en la investigación, con la finalidad de darles uniformidad y constancia en la validación de la información obtenida. Analizando las funciones que cumplen las hipótesis en una investigación, no hay duda acerca del papel importante que éstas desempeñan en un estudio y en el campo científico en general.

7.7.4 Clases de hipótesis

Algunos de los tipos de hipótesis más usuales en la investigación son los siguientes:

Hipótesis de trabajo Hipótesis inicial que plantea el investigador al dar una respuesta anticipada al problema objeto de investigación. Un ejemplo de esta hipótesis podría ser:

> Supongamos que existe interés por analizar el problema del desempleo en una determinada ciudad del país y el investigador se propone la siguiente hipótesis:
>
> H_A: las principales causas del desempleo en la ciudad están determinadas por las medidas económicas del gobierno nacional.
>
> H_A: los hijos adolescentes de padres separados tienen mayor riesgo de incurrir en el consumo de licores que los hijos adolescentes que viven con sus dos padres en familia nuclear.

Hipótesis nula Hipótesis que indica que la información que se va a obtener es contraria a la hipótesis de trabajo. En los anteriores ejemplos, las hipótesis nulas serían:

> H_o: el fenómeno del desempleo en la ciudad no está determinado por las medidas económicas del gobierno nacional.
>
> H_o: no hay diferencia significativa en el riesgo de consumo de licores por los hijos adolescentes de padres separados y de los padres casados que viven juntos en familia nuclear.

Hipótesis descriptivas Hipótesis o suposiciones respecto a rasgos, características o aspectos de un fenómeno, un hecho, una situación, una persona, una organización, etcétera.

> H_i: las principales características del desempleo en la ciudad son la edad, el nivel educativo y el sexo.
>
> H_i: los rasgos propios de un adolescente que manifieste conductas suicidas son...

www.FreeLibros.me

Hipótesis estadísticas Hipótesis o suposiciones formuladas en términos estadísticos.

H_1: 25% de la población desempleada en la ciudad corresponde a personas con nivel académico profesional.

H_1: en el 87% de las familias donde la cabeza de familia padece desempleo, cada uno de los miembros de tales familias presentan altos grados de estrés.

7.7.5 Procedimiento para verificar hipótesis

Uno de los aspectos importantes relacionados con la hipótesis es el procedimiento estadístico que debe seguirse para verificar o realizar una prueba de hipótesis.

En general, la prueba de hipótesis puede sintetizarse en los siguientes pasos:

1. *Formular la hipótesis*: consiste en plantear la hipótesis nula (H_o) y la hipótesis alterna o de trabajo (H_A) del problema objeto de la investigación.
2. *Elegir la prueba estadística adecuada*: el investigador elige la prueba estadística teniendo en cuenta las características del tema de investigación. (En el numeral 7.11 se describe el procesamiento de la información).
3. *Definir el nivel de significación*: para la prueba de hipótesis es necesario definir un porcentaje o nivel de confianza dentro del cual se aceptará o rechazará la hipótesis. Es usual usar valores para $\alpha = 0{,}01$, $\alpha = 0{,}05$ ó $\alpha = 0{,}10$.
4. *Recolectar los datos de una muestra representativa*: consiste en obtener la información de la población o muestra objeto del estudio.
5. *Estimar la desviación estándar de la distribución muestral de la media*: se utiliza la siguiente fórmula:

$$S_x = \frac{S}{\sqrt{n}}$$

Donde:

S_x: desviación estándar de la distribución muestral de la media.

S: desviación estándar de la muestra.

n: tamaño de la muestra.

6. *Transformar la media de la muestra en valores z o t, según la prueba estadística seleccionada.*
7. *Tomar la decisión estadística*: Consiste en comparar el valor de *z* o *t* calculado en el paso anterior, con el respectivo valor de z o t crítico (valor en tabla), según el nivel de significación elegido en el numeral 3.
8. *Conclusión*: Consiste en llegar a una conclusión de rechazo o aceptación de la hipótesis objeto del estudio.

www.FreeLibros.me

En la sección correspondiente al procesamiento de información (7.11) el lector podrá encontrar algunos ejemplos de prueba de hipótesis aplicados al campo de las ciencias económicas y administrativas.

7.7.5 Hipótesis y variables

Para probar las hipótesis es necesario identificar el concepto de variable, porque las hipótesis son suposiciones acerca de variables. Pero ¿qué es una variable? De acuerdo con Rojas Soriano (1981), una variable "es una característica, atributo, propiedad o cualidad que puede estar o no presente en los individuos, grupos o sociedades; puede presentarse en matices o modalidades diferentes o en grados, magnitudes o medidas distintas a lo largo de un *continuum*" (p. 87) .

En este sentido, una hipótesis es una suposición de la relación entre características, atributos, propiedades o cualidades que definen el problema objeto de la investigación. Estas características o propiedades se definen como variables de investigación.

7.7.6 Tipos de variables

En las hipótesis causales, es decir, aquellas que plantean relación entre efectos y causas, se identifican tres tipos de variables: *independientes, dependientes* e *intervinientes*. Estos mismos tipos de variables pueden estar presentes en las hipótesis correlacionales cuando se explica la correlación.

Independiente Se denomina *variable independiente* a todo aquel aspecto, hecho, situación, rasgo, etcétera, que se considera como la "causa de" en una relación entre variables.

Dependiente Se conoce como *variable dependiente* al "resultado" o "efecto" producido por la acción de la variable independiente.

Interviniente Las *variables intervinientes* son todos aquellos aspectos, hechos y situaciones del medio ambiente, las características del sujeto/objeto de la investigación, el método de investigación, etcétera, que están presentes o "intervienen" (de manera positiva o negativa) en el proceso de la interrelación de las variables independiente y dependiente.

En toda investigación experimental es muy importante identificar y hacer un adecuado control de variables intervinientes para tener confiabilidad sobre la interdependencia de las variables independiente y dependiente, debido a que en un evento puede presentarse el caso de que las variables intervinientes alteren la relación entre estas variables. Por ejemplo, en un estudio en que se pretende medir el efecto sobre el volumen de ventas de un programa de capacitación en estrategias de mercadeo, realizado al personal del área de ventas de una determinada empresa, si no existe control de variables intervinientes será difícil afirmar que, en caso de presentarse variación en el volumen de ventas, esa variación se deba al programa de capacitación recibido por el personal, ya que la variación podría obedecer a otras variables como la temporada, la baja en los precios del producto, etcétera, y no a la capacitación del personal.

www.FreeLibros.me

El control de variables intervinientes se realiza indicando cuáles variables serán controladas, cómo será el control y por qué se controla.

Las principales variables que se controlan son las que tienen que ver con:

- *Condiciones ambientales*: se realiza el estudio en condiciones constantes o se controlan los posibles cambios; ejemplos: empresas en entornos económicos similares; personas en condiciones similares.
- *Experiencia de los sujetos (población o muestra) del estudio*: se seleccionan sujetos o se selecciona una población con similar experiencia en la variable objeto de medición; ejemplo: empresas o personas que tienen experiencias semejantes.
- *Homogeneidad en las características de la población objeto del estudio*: se selecciona una población con características similares; ejemplo: empresas que tienen características similares (tamaño, sector económico, estilo de dirección, posicionamiento en el mercado, etcétera); personas que poseen características similares (edad, grado académico, estrato socioeconómico, etcétera).

En el ejemplo 7.22 se ilustran los tres tipos de variables en una hipótesis causal.

EJEMPLO 7.22 **Tipos de variables en una hipótesis causal**

En el campo de los negocios

Ejemplo 1

Hipótesis 1: el nivel de productividad del personal de una organización está determinado por el grado de capacitación académica que tiene cada persona.

Variable independiente: grado de capacitación académica (causa).

Variable dependiente: nivel de productividad (efecto).

Variables intervinientes: ambiente laboral, temporada del año, nivel salarial, estilo de dirección, rasgos de personalidad.

Ejemplo 2

Hipótesis 2: el costo del dinero (tasa de interés) en el mercado determina el monto de inversión de las empresas.

Variable independiente: costo del dinero (tasa de interés).

Variable dependiente: volumen o monto de inversión por parte de las empresas.

Variables intervinientes: condiciones sociales y económicas del país, capacidad de inversión por parte de la empresa.

En psicología y educación

Ejemplo 3

Hipótesis 3: el grado de efectividad de una psicoterapia familiar depende del tipo de estrategia que utilice el equipo psicoterapeuta.

www.FreeLibros.me

Variable independiente: tipo de estrategia utilizado por el equipo psicoterapeuta.

Variable dependiente: grado de efectividad de la psicoterapia.

Variables intervinientes: experiencia de los psicoterapeutas, condiciones sociales de la familia intervenida, experiencias previas de la familia en intervenciones psicoterapéuticas, condiciones ambientales en las que se realiza la psicoterapia, etcétera.

Ejemplo 4

Hipótesis 4: el grado de calidad del aprendizaje de un tema, por un estudiante, depende principalmente del grado de interés que éste tenga del mismo.

Variable independiente: grado de interés que un estudiante tiene sobre un tema.

Variable dependiente: calidad del aprendizaje de un tema.

Variables intervinientes: estado de ánimo del estudiante, condiciones ambientales en donde se dan el aprendizaje, las estrategias pedagógicas utilizadas durante el estudio, etcétera.

7.7.7 Conceptuación y operacionalización de las variables

Una vez identificadas las variables objeto del estudio, es necesario conceptuarlas y operacionalizarlas.

Conceptuar una variable quiere decir definirla, para clarificar qué se entiende por ella.

Operacionalizar una variable significa traducir la variable a indicadores, es decir, traducir los conceptos hipotéticos a unidades de medición.

EJEMPLO 7.23 Conceptuación y operacionalización de variables

En el campo de la economía y los negocios

Ejemplo 1

Cuando se habla de la variable *ingreso*, conceptuar la variable significa especificar qué se entiende por ingreso.

Cuando se hace referencia a la variable *capacitación*, conceptuar la variable significa definir el término capacitación.

Ejemplo 2

- La variable *ingreso* se operacionaliza cuando se define en una unidad de medición, como puede ser cantidad de ingresos ($) en alguna moneda.
- La variable *capacitación* se operacionaliza cuando se traduce a número de años de formación académica.
- La variable *pobreza* se operacionaliza cuando se traduce a promedio de ingresos de las personas.

www.FreeLibros.me

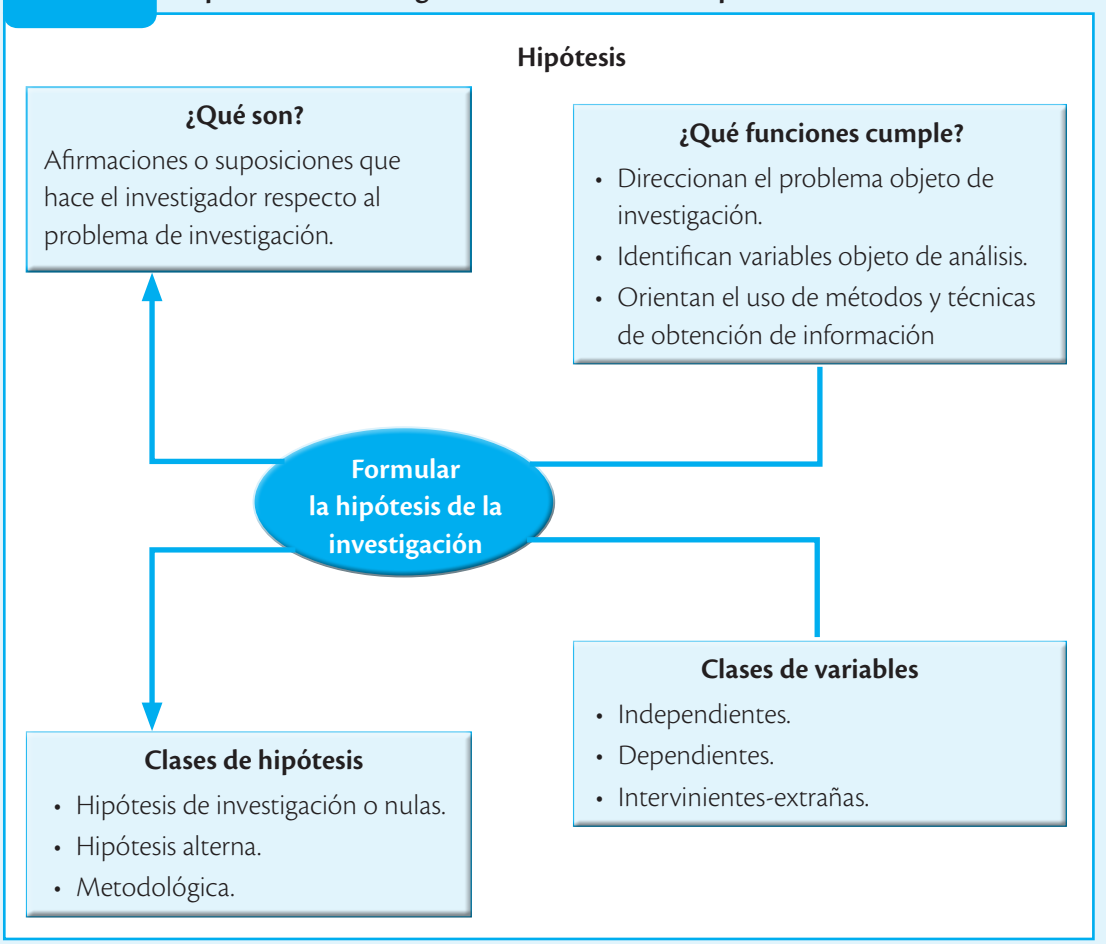
En psicología y educación

La variable *ansiedad* se operacionaliza cuando se define en términos de alguna medida como grado de ansiedad, y se mide por manifestaciones como cambio en las condiciones psicogalvánicas de la piel, en la temperatura del cuerpo, en el ritmo cardiaco, en la sudoración del cuerpo, etcétera.

La variable *calidad del aprendizaje* se operacionaliza cuando se define alguna forma de medir el aprendizaje. Es la cantidad y calidad del conocimiento que un estudiante tiene sobre un determinado tema.

La Figura 7.11 sintetiza el concepto y las funciones de las hipótesis en un proceso de investigación, así como los principales tipos de éstas y las variables que tienen relación con las mismas.

FIGURA 7.11 **El proceso de investigación. Formulación de hipótesis**



www.FreeLibros.me

RESUMEN

Se formulan hipótesis cuando en la investigación se quiere probar una suposición y no sólo mostrar los rasgos característicos de una determinada situación. Es decir, se formulan hipótesis en las investigaciones que buscan probar el impacto que tienen algunas variables entre sí, o el efecto de un rasgo o una variable en relación con otro(a). Las investigaciones de tipo descriptivo no requieren formular hipótesis; es suficiente plantear algunas preguntas de investigación que, como ya se anotó, surgen del planteamiento del problema, de los objetivos y, por supuesto, del marco teórico que soporta el estudio.

En general, todo proyecto de investigación requiere preguntas de investigación, pero sólo aquellos que buscan evaluar relación entre variables o explicar causas requieren la formulación de hipótesis. En el caso de la investigación experimental, siempre es necesario partir de hipótesis, las cuales guiarán el respectivo estudio.

Una hipótesis es una suposición o solución anticipada al problema objeto de la investigación y, por tanto, la tarea del investigador debe orientarse a probar tal suposición o hipótesis. Aceptar una hipótesis como cierta no implica concluir respecto de la veracidad de los resultados obtenidos, sino que sólo se aporta evidencia en su favor.

Algunos de los tipos de hipótesis más usuales en la investigación son los siguientes:

- *Hipótesis de trabajo*: la hipótesis inicial que plantea el investigador al dar una respuesta anticipada al problema objeto de investigación.
- *Hipótesis nula*: una hipótesis que indica que la información por obtener es contraria a la hipótesis de trabajo.
- *Hipótesis descriptivas*: aquellas hipótesis o suposiciones respecto a rasgos, características o aspectos de un fenómeno, un hecho, una situación, una persona, una organización, etcétera.
- *Hipótesis estadísticas*: hipótesis o suposiciones formuladas en términos estadísticos.

En general, la prueba de hipótesis puede sintetizarse en los siguientes pasos:

1. Formulación de la hipótesis.
2. Elección de la prueba estadística adecuada.
3. Definición del nivel de significación.
4. Recolección de los datos de una muestra representativa.
5. Estimación de la desviación estándar de la distribución muestral de la media.

www.FreeLibros.me

6. Transformación de la media de la muestra en valores *z* o *t*, según la prueba estadística seleccionada.
7. Toma de la decisión estadística.
8. Conclusión.

Para probar las hipótesis es necesario identificar el concepto de variable que, según Rojas Soriano (1981), "es una característica, atributo, propiedad o cualidad que puede estar o no presente en los individuos, grupos o sociedades; puede presentarse en matices o modalidades diferentes o en grados, magnitudes o medidas distintas a lo largo de un ***continuum***" (p. 87). En la prueba de hipótesis suelen identificarse tres tipos de variables: *independientes*, *dependientes* e *intervinientes*.

Se denominan *variables independientes* a todos aquellos aspectos, hechos, situaciones, rasgos, etcétera, que se consideran como la "causa de" en una relación entre variables; *variables dependientes*, al "resultado" o "efecto" producido por la acción de la *variable independiente*, y *variables intervinientes*, todos aquellos aspectos, hechos y situaciones del medio ambiente, las características del sujeto/objeto de la investigación, el método de investigación, etcétera, que están presentes o "intervienen" (de manera positiva o negativa) en el proceso de la interrelación de las variables independiente y dependiente.

EJERCICIOS DE REPASO Y ANÁLISIS

1. ¿Qué son las hipótesis de investigación?
2. ¿Por qué son importantes las hipótesis en investigación y cómo se verifican?
3. Explique la relación entre problema de investigación, objetivos, tipo de investigación e hipótesis.
4. Seleccione un tema de investigación, formule los objetivos, defina el tipo de investigación y evalúe la pertinencia o no de formular hipótesis. (En caso de ser necesario, formule la hipótesis de investigación).
5. Muestre por lo menos tres ejemplos de hipótesis de investigación para tres temas de interés en el campo de su disciplina e identifique las variables de las respectivas hipótesis, con sus definiciones (conceptual y operacional).
6. Consulte en la biblioteca y revise por lo menos tres estudios de investigación reciente en los que se hayan planteado hipótesis de investigación. Obtenga información relacionada con el título, los autores, los objetivos de la investigación, las hipótesis, la prueba de hipótesis y las conclusiones.
7. En un taller de clase, analicen la información del desarrollo de los puntos 1 a 6.

www.FreeLibros.me

7.8 Diseño experimental de la investigación. Definición

La definición de un diseño de investigación está determinada por el tipo de investigación que va a realizarse y por la hipótesis que va a probarse durante el desarrollo de la investigación. Se habla de diseños cuando está haciéndose referencia a la investigación experimental, que consiste en demostrar que la modificación de una variable (independiente) ocasiona un cambio predecible en otra (variable dependiente).

La investigación experimental se realiza mediante los llamados diseños, que son un conjunto de procedimientos con los cuales se manipulan una o más variables independientes y se mide su efecto sobre una o más variables dependientes.

7.8.1 Diseños experimentales

En la investigación experimental existen diversos tipos de diseño, que se clasifican de diferentes formas. Sin embargo, la clasificación más usada, según Salkind (1998) e investigadores como Briones (1985), es la de Campbell y Stanley, quienes identifican tres categorías generales de diseños de investigación: preexperimentales, cuasi experimentales y experimentales verdaderos. Para Salkind (1998), esta clasificación se fundamenta en dos características básicas de los diseños: 1) el grado de control que se ejerce sobre las variables objeto de estudio, y 2) el grado de aleatoriedad con que se asignan los sujetos de la investigación a un grupo o a varios de ellos.

Diseños experimentales verdaderos

Se caracterizan por un alto grado de control de las variables y porque efectúan asignación aleatoria de los sujetos a los grupos (experimental y control) participantes en la investigación.

De acuerdo con Salkind (1998), para que en investigación un diseño sea un experimento verdadero, se requiere que haya una manipulación intencional de una o más variables independientes, que se realice asignación aleatoria de los sujetos participantes en la investigación a cada uno de los grupos (experimental y de control), y que se ejerza un riguroso control sobre las variables objeto de medición y sobre las variables extrañas que pueden tener algún impacto en los resultados del experimento.

Estas variables son específicas de cada caso y pueden ser el tipo de prueba que se aplique (debe ser la misma para todos), las condiciones en que se aplica (las mismas para todos), la experiencia previa de los sujetos de la investigación (que sea homogénea), etcétera.

En el campo de la administración, un ejemplo clásico de un diseño experimental es el experimento Hawthorne, dirigido por Elton Mayo entre 1927 y 1929, en la fábrica de Western Electric Company, situada en Chicago, en el barrio Hawthorne, de donde tomó su nombre.

www.FreeLibros.me

Existen diversos tipos de diseños experimentales verdaderos como los siguientes:

- Diseño de medición previa y posterior con grupo de control.
- Diseño de Solomon para cuatro grupos.
- Diseño con medición posterior y grupo de control.
- Diseño de series cronológicas.
- Diseños factoriales.

Al finalizar esta sección, en el ejemplo 7.24, se muestra la aplicación de estos diseños al campo de las ciencias económicas y administrativas.

Diseños cuasiexperimentales

Los diseños cuasiexperimentales se diferencian de los experimentales verdaderos porque en aquéllos el investigador ejerce poco o ningún control sobre las variables extrañas, los sujetos participantes de la investigación se pueden asignar aleatoriamente a los grupos y algunas veces se tiene grupo de control.

Estos diseños usualmente se utilizan para grupos ya constituidos.

Los siguientes son algunos diseños cuasiexperimentales:

- Diseños de un grupo con medición antes y después.
- Diseños con grupo de comparación equivalente.
- Diseños con series de tiempos interrumpidos.

Diseños preexperimentales

Presentan el más bajo control de variables y no efectúan asignación aleatoria de los sujetos al experimento, y son aquellos en los que el investigador no ejerce ningún control sobre las variables extrañas o intervinientes, no hay asignación aleatoria de los sujetos participantes de la investigación ni hay grupo control.

Algunos diseños preexperimentales son:

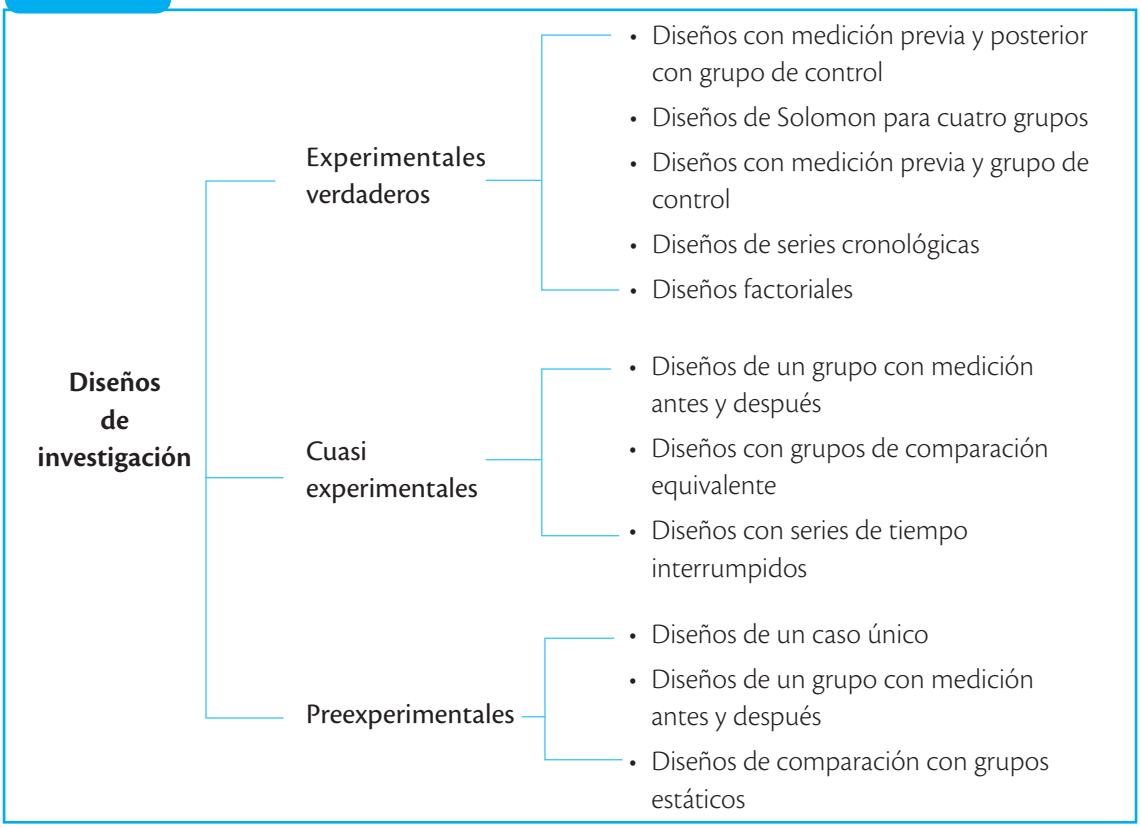
- Diseño de un caso único.
- Diseño de un grupo con medición (prueba) previa y posterior.
- Diseño de comparación con un grupo estático.

Algunos de estos diseños aplicados se encuentran en el ejemplo 7.24, al final de esta sección.

La Tabla 7.2 sintetiza los principales tipos de diseños experimentales en investigación científica.

www.FreeLibros.me

TABLA 7.2 **Tipos de diseño experimentales**



7.8.2 Validez de los experimentos

Un aspecto fundamental en el diseño de experimentos es la *validez* de los resultados que se obtengan de éstos. La pregunta básica en todo experimento de investigación es ¿son válidos los resultados?

En un diseño experimental, cualquier variable extraña que interfiera con la capacidad de efectuar inferencias causales se considera una amenaza para la validez de sus resultados; por tanto, en investigación es importante el control de variables extrañas.

Según Malhotra (1997), en la investigación experimental se presentan dos clases fundamentales de validez, a las que hay que atender: la interna y la externa.

Validez interna

La *validez interna* mide si la acción de las variables independientes o los tratamientos producen los efectos en la variable dependiente. Si el investigador demuestra que la variable experimental (variable independiente) o de tratamiento produjo las diferencias observadas en la variable dependiente, se dice que el experimento tiene validez interna.

www.FreeLibros.me

Amenazas contra la validez interna

Historia Se refiere a cualquier evento o circunstancia distinto de los manipulados por el investigador (variable independiente), presente durante el experimento y que puede afectar el resultado de la variable dependiente; por ejemplo, los cambios en las condiciones ambientales que se presenten durante el tiempo de duración del experimento.

Maduración Se refiere a los cambios en el interior de las unidades de prueba que ocurren durante el tiempo de realización del experimento; por ejemplo, la experiencia que pudieran ir adquiriendo los sujetos participantes en la investigación.

Variación en los instrumentos Se refiere a cualquier cambio en los instrumentos de recolección de la información que afecte las mediciones obtenidas; por ejemplo, utilizar instrumentos de medición diferentes o modificarlos durante la investigación.

Selección sesgada Se presenta cuando se realiza una asignación inadecuada de las unidades de prueba a las condiciones del tratamiento; por ejemplo, seleccionar un grupo de control con características diferentes de las del grupo experimental.

Mortalidad Consiste en la pérdida de unidades de prueba o sujetos participantes en la investigación; por ejemplo, participantes en una investigación que se retiran durante la realización del experimento.

Validez externa

A diferencia de la validez interna, que indica si la variable independiente o acción realizada sobre un grupo o sujeto experimental, X, influyó realmente en las mediciones observadas, la *validez externa* se centra en la posibilidad de que los resultados del experimento se generalicen a personas, medios y tiempos en el ambiente real; por ejemplo, generalizar los resultados de un experimento realizado con un grupo de empresarios de los sectores de las pequeñas y medianas empresas (pymes), que participaron en un curso de capacitación para mejorar la competitividad de ese sector en el comercio exterior, a todas las empresas del sector.

Amenazas contra la validez externa

Igual que en la validez interna, es decir, en la capacidad de un experimento para demostrar que una(s) variable(s) independiente(s) (VI) fue(ron) la(s) que produjo(eron) los cambios en la(s) variable(s) dependiente(s) (VD), hay amenazas que afectan esa validez y, por tanto, es necesario identificar y controlar estas amenazas; en la validez externa, la cual se refiere al grado en que los resultados de un experimento se pueden generalizar desde una muestra específica a una población general (Buendía, et al., 2001) o de si los resultados de un experimento se lograrán o no generalizar a otro entorno (Salkind, 1998).También se presentan amenazas que es necesario conocer y controlar de parte del investigador para su control y de parte de quienes usan los resultados de tales investigaciones.

A continuación se presentan las principales amenazas contra la validez externa de los experimentos.

www.FreeLibros.me

Interferencia de múltiples tratamientos Se presenta cuando durante un experimento se produce de manera simultánea al tratamiento experimental otro tratamiento a los participantes del experimento, que afecta los resultados del experimento. Un ejemplo de esta amenaza ocurre cuando en un experimento, cuya finalidad es probar un determinado método pedagógico, para mejorar el rendimiento académico de los estudiantes, se cambian las condiciones ambientales del lugar en donde se desarrolla el experimento o se ofrece algún incentivo por el buen rendimiento.

Grado de representatividad de la muestra o población La validez de un experimento depende, en alto grado, de la representatividad de la muestra de participantes con los que se lleva a cabo el experimento. Por ello, es necesario que la muestra con que se realice el experimento sea sistemáticamente estimada y seleccionada para que haya representatividad de sus resultados y los mismos puedan generalizarse, tanto a otras poblaciones como a otros contextos, con las debidas adaptaciones. Un ejemplo de amenaza por bajo grado de representatividad de la muestra ocurre cuando un experimento se efectúa con muestras que no han sido estimadas ni seleccionadas con criterios estadísticos adecuados.

Efectos del experimentador Se presenta cuando en un experimento el investigador adopta un determinado comportamiento, que es particular y no tendrá lugar con otras poblaciones o en otros contextos. Por ejemplo, cuando en un experimento el investigador asume un comportamiento, ya sea afectivo o apático, con los participantes del experimento por circunstancias específicas, pero ese comportamiento no será el cotidiano en otros contextos.

Efecto Hawthorne Se presenta tal amenaza cuando los participantes de la experimentación cambian de comportamiento ante la presencia del experimentador. Para ejemplificar este efecto, es posible mencionar la actitud que pueden asumir los sujetos de un experimento, el cual busca aumentar la productividad de un grupo de trabajadores, y las personas con las cuales se realiza el experimento modifican su comportamiento ante la presencia del experimentador por el solo hecho de ser observados o por la percepción que éstos lleguen a crear por las implicaciones que se derivan de los resultados del estudio.

Efecto por la novedad y las interrupciones cotidianas Según Buendía E. et al. (2001) los resultados de un experimento llegan a afectarse porque el experimento puede convertirse en un espacio de novedad o de diversidad, que permite salir de la rutina diaria a los participantes del experimento y los efectos del tratamiento tienden a ser más secundarios que determinantes de los mismos. Un ejemplo sería cuando personas que cotidianamente están en ambientes o condiciones rutinarias, al formar parte de un experimento, tienden a considerar el mismo como un espacio para romper la monotonía o rutina cotidiana, lo cual necesariamente afecta los resultados del estudio.

Efecto por sensibilización de pretest o postest Otra amenaza externa de los experimentos son los efectos que se pueden generar en los resultados de un experimento, como consecuencia de la aplicación de pruebas previas o posteriores a un experimento.

www.FreeLibros.me

Las pruebas previas llegan a sensibilizar algún grado de aprendizaje, que puede influir en los resultados del experimento. En el caso de las pruebas posteriores, cuando éstas actúan como estrategias de aprendizaje, producen similares efectos que las pruebas previas, con incidencia directa en los resultados del estudio. Un ejemplo de tal variable es el caso de un experimento sobre aprendizaje de un determinado tema: cuando se hace un pretest, los resultados del mismo afectan la actitud de las personas, ya sea positiva o negativamente, y, por tanto, el estudio.

Control de variables

Para efectos del tema de la investigación experimental, la palabra "control", como señala McGuigan (1996) "indica que el experimentador tiene cierto poder sobre las condiciones de un experimento; poder que consiste en manejar de manera sistemática las diferentes variables, tanto independientes como extrañas, en un esfuerzo por llegar a la debida conclusión empírica" (p. 60), es decir, llegar a conocer la verdadera relación entre la(s) variable(s) independiente(s) y la(s) dependiente(s).

Control de la variable independiente De acuerdo con McGuigan (1996), en este caso el término "control" se aplica cuando la variable independiente se manipula o modifica de manera intencional, conocida y específica por parte del investigador.

Control de variables extrañas Las variables extrañas son variables diferentes a la variable independiente, que si no se controlan actúan durante el desarrollo de un experimento y afectan a la variable dependiente, generando dificultad para conocer la relación entre la variable independiente y la variable dependiente, que es el objetivo básico de los experimentos. En este sentido, control significa regular aquellas variables que pueden tener relación directa y, por tanto, afectar los resultados de un experimento.

Algunos ejemplos de control de variables extrañas en los experimentos son: la definición de la edad, el sexo, el estado civil, el nivel académico, la experiencia previa con la variable de experimentación de los participantes, las condiciones ambientales donde se realizan los experimentos, la hora, etcétera. Tales variables deben ser específicas en cada experimento, es decir, las variables extrañas en un experimento con una determinada finalidad son diferentes a las de otro experimento con otro propósito.

Por ejemplo, en un experimento donde se quiere probar una estrategia pedagógica para mejorar el grado de aprendizaje de un determinado conocimiento, algunas variables extrañas por controlar son: la edad, el grado académico, la experiencia previa de los participantes en la estrategia pedagógica, la hora, las condiciones ambientales en las que se realiza y las características personales de los experimentadores.

En otro experimento cuya finalidad es probar la eficacia de una nueva estrategia, para dejar la conducta del consumo de cigarrillo, algunas de las variables extrañas que debe controlar el investigador son: la frecuencia del hábito del consumo de cigarrillo, las experiencias previas de los participantes en la terapia para dejar el hábito del consumo del cigarrillo,

www.FreeLibros.me

el tiempo que lleva la persona con este hábito de fumar, el sexo del fumador, el nivel académico del fumador, el ambiente social que comparte, etcétera.

En general, un control de variables en un experimento consiste en identificar aquellas variables que pueden afectar durante el desarrollo del experimento los resultados del mismo al generarse dificultad para poder probar que ha sido realmente la variable independiente la que ha generado los respectivos resultados y no otras variables. Además, es importante señalar que del grado de control de estas variables depende el grado de validez tanto interna como externa de los experimentos, sin embargo, recuérdese que en los experimentos aunque es complejo el control total de las variables extrañas, sí es necesario el máximo control de las mismas.

Ahora, ¿cómo se lleva un control de variables extrañas en un experimento? Hay varios métodos básicos para controlar variables extrañas; algunos son:

- *La igualación*: consiste en la asignación aleatoria de los participantes del experimento a los grupos con que se realizará el estudio.
- *La constancia de las condiciones*: consiste en mantener constantes las condiciones en las cuales se realiza el experimento durante el tiempo que dure éste.
- *La eliminación*: consiste en utilizar condiciones de laboratorio para el desarrollo de los experimentos, porque tales condiciones son las más adecuadas para eliminar variables como ruidos, mantener constantes las temperaturas, aislar a los participantes de situaciones que afectarían el estudio, etcétera.
- *Presencia de grupos de control*: consiste en agrupar los sujetos de la investigación en subgrupos, con el propósito de aplicar el tratamiento a la mitad de los subgrupos y la otra mitad; aunque se mantiene en iguales condiciones que a los grupos del tratamiento, no se actúa sobre ellos, pero sobre los cuales, al igual que a los demás grupos, se les mide la variable dependiente del experimento, con el propósito de comparar los resultados de los diferentes grupos. A los grupos que no se les aplica tratamiento, se les denomina grupos de control; reciben este nombre porque su función es servir de control de variables extrañas, permitiendo comparar resultados con grupos similares que sí reciben el efecto de variables independientes.

7.8.3 Notación convencional de los experimentos

Para el estudio de los experimentos se utiliza un sistema estándar de notación:

X : indica la variable independiente (acción realizada sobre la población objeto del estudio); también se le conoce como tratamiento.
O : se emplea para indicar la medición de la variable dependiente.
R : se utiliza para indicar la asignación aleatoria de las unidades de prueba a los grupos en el experimento.
G : indica el grupo objeto del estudio.
GE : grupo experimental.
GC : grupo control.

www.FreeLibros.me

La Tabla 7.3 muestra la relación de los diferentes tipos de diseños experimentales con la aleatoriedad de los sujetos objeto de la investigación, el control de variables extrañas y la presencia o no de grupos de control.

TABLA 7.3 **Diseños de investigación**

	Experimentales	Cuasiexperimentales	Preexperimentales
Asignación de los sujetos	Aleatorios	Aleatorios voluntarios	Voluntarios
Control de variables	Máximo	Baja	Ninguno
Grupo de control	Siempre	Algunas veces	Ninguno

La Figura 7.12 muestra la relación del diseño experimental de investigación con los otros componentes según el diagrama de espiral del proceso de investigación propuesto en este libro.

FIGURA 7.12 **El proceso de investigación. Definición del diseño**

www.FreeLibros.me

EJEMPLO 7.24 **Tipos de diseños experimentales de investigación**

A continuación se ilustran algunos diseños de investigación experimental que sirven de guía para realizar futuras investigaciones, en las cuales se requiere el uso de estos diseños.

Un investigador del campo de las ciencias económicas y administrativas está interesado en conocer el impacto de un programa de capacitación referente a "Productividad, competitividad y comercio exterior" para el sector de las pequeñas y medianas empresas (pymes) del país, en el volumen de las exportaciones del sector.

La hipótesis de investigación propuesta por el investigador es:

H_o : un programa de capacitación en el tema "productividad, competitividad y comercio exterior", desarrollado con directivos y empresarios del sector de las pymes, conduce a un aumento en el volumen de exportaciones de las empresas del sector.

Para realizar esta investigación se propusieron siete diseños que se describen a continuación.

A. Diseño preexperimental

Este diseño se caracteriza por el *nulo* o *poco control* de variables extrañas durante la realización del experimento.

Diseño con un solo grupo

Este diseño preexperimental tiene las siguientes características:

Es un diseño sin grupo control y donde sólo se efectúa una medición posterior.

Esquema del diseño: $G \quad X \quad O_1$

Donde:

G : grupo objeto de estudio (empresas del sector de las pymes).

X : tratamiento o variable independiente (programa de capacitación).

O_1 : medición de la variable dependiente (variación en el volumen de las exportaciones en un período de un año, después de haber tomado el curso).

El diseño consiste, entonces, en tomar un grupo voluntario de empresas, de las cuales sus empresarios o directivos participarán en el programa de capacitación; se desarrolla la capacitación y, luego, se realiza la medición del volumen de exportaciones para evaluar el impacto del curso en dicha variable.

Análisis del diseño. Con este diseño no se garantiza que los resultados que se obtengan en una variación en el volumen de las exportaciones obedezca al efecto causado por la capacitación recibida en el curso, debido a que no se controlan situaciones del medio que pueden ser las causales de las variaciones (si las hay).

www.FreeLibros.me

Algunas de estas situaciones o variables extrañas que pueden afectar los resultados del experimento son los cambios en las condiciones económicas o políticas del país.

B. Diseños cuasiexperimentales

Diseño de comparación con dos grupos estáticos

Este diseño presenta las siguientes características:

Es un diseño que utiliza un *grupo experimental* y uno de control. Sin embargo, los sujetos o las unidades de prueba no se asignan de manera aleatoria a ningún grupo ni se realizan mediciones previas al experimento de la variable dependiente.

Esquema del diseño:

Grupo experimental	X	O_1
Grupo control	-	O_2

Donde:

X : variable independiente (programa de capacitación).

O_1 : medición de la variable dependiente (variación en el volumen de las exportaciones).

O_2 : medición de la variación del volumen de exportaciones del grupo control durante el tiempo de realización del experimento (este grupo no recibe capacitación).

Este diseño consiste en disponer de dos grupos voluntarios: uno de ellos participará en el programa de capacitación (este grupo se denomina grupo experimental), mientras el otro grupo no recibirá ninguna capacitación, pero servirá de grupo control.

Análisis del diseño. A pesar de que en este diseño se utilizan dos grupos de empresarios de los cuales uno recibe capacitación y el otro no, el hecho de que los empresarios o directivos no se asignen aleatoriamente a los grupos, la comparación que se haga de los resultados obtenidos al final del experimento y se obtenga diferencia positiva del grupo que recibe capacitación respecto al que no la recibe, no garantiza que esa diferencia se deba a la capacitación, puesto que en el caso anterior no se controlan variables del medio que podrían afectar los resultados del experimento.

Diseño de un grupo con medición antes y después

Este diseño tiene las siguientes características:

Es un diseño de un solo grupo con medición previa (antes) y posterior (después) de la variable dependiente, pero sin grupo control.

Esquema del diseño: $G \quad O_1 \quad X \quad O_2$

Donde:

X : variable independiente (programa de capacitación).

O_1 : medición previa (antes del curso) de la variable dependiente (volumen de exportaciones).

O_2 : medición posterior (después de tomar el curso) de la variable dependiente.

www.FreeLibros.me

El diseño consiste en contar con un grupo voluntario de empresas cuyos empresarios o directivos participarán en el programa de capacitación. A ese grupo de empresas se le hará una evaluación previa sobre la variable objeto de estudio (volumen de exportaciones) y, después de haber recibido la capacitación, se realizará una nueva medición con el propósito de comparar los resultados antes y después del curso.

Análisis del diseño. La evaluación del grado de conocimiento en competitividad y comercio exterior en los empresarios y directivos empresariales, realizada antes y después de la capacitación, sirve de parámetro para medir el efecto del curso en la variación de las exportaciones; pero como no se controlan variables extrañas que puedan afectar a los empresarios en sus actividades de comercio exterior, entonces es posible concluir que los cambios que se presenten en las exportaciones obedezcan realmente al curso.

C. Diseños experimentales verdaderos

Para superar las deficiencias de los anteriores diseños, el investigador ha propuesto cuatro diseños experimentales que permiten contar con resultados sobre los cuales pueden hacerse inferencias.

Diseño con dos grupos aleatorios

Este diseño presenta las siguientes características:

Es un diseño con un grupo experimental y uno de control, que incluye la asignación aleatoria de los sujetos o las unidades de análisis a ambos grupos. Sin embargo, no se efectúa medición previa de la variable dependiente a ningún grupo.

Ofrece un alto control de variables extrañas.

El esquema del diseño es:	Grupo experimental	(R)	X	O_1
	Grupo control	(R)	—	O_2

Donde:

R : asignación aleatoria de las unidades objeto de estudio a los grupos.

X : variable independiente (programa de capacitación).

O_1 : medición de la variable dependiente (variación del volumen de las exportaciones) en el grupo experimental.

O_2 : medición de la variación objeto de estudio en el grupo control.

En este diseño, el investigador cuenta con un número determinado de pequeñas y medianas empresas (pymes) interesadas en el comercio exterior, que tienen experiencia de por lo menos un año en el campo de las exportaciones. Además, durante los dos últimos años no han recibido capacitación en el tema de la competitividad ni en el del comercio exterior.

Las empresas se asignan aleatoriamente a dos grupos; uno de éstos participará en el programa de capacitación sobre "Productividad, competitividad y comercio exterior", mientras que el otro no recibirá capacitación en el tema.

www.FreeLibros.me

Después de haber realizado la capacitación al grupo experimental, se mide la variable objeto de estudio y se comparan los resultados de los dos grupos (experimental y control). En el caso de presentarse diferencias significativas en la comparación de resultados, se infiere que el programa de capacitación ha tenido efecto directo sobre la variable dependiente (exportaciones).

Análisis del diseño. En este diseño, el hecho de que los empresarios o directivos se asignen aleatoriamente a los dos grupos hace que estos grupos sean relativamente idénticos entre sí, y teniendo en cuenta que un grupo recibe capacitación mientras que el otro no, entonces la diferencia de los resultados que se obtenga al comparar los dos grupos, puede afirmarse, con relativa certeza, que obedece a efectos del programa de capacitación.

Diseño con medición previa y posterior con grupo de control

Este diseño presenta las siguientes características:

Es un diseño que incluye la *asignación aleatoria* de los sujetos o las unidades de análisis, tanto al grupo experimental como al grupo control, y se realiza medición previa y posterior de la variable dependiente a ambos grupos.

Esquema del diseño:	Grupo experimental	(R)	O_1	X	O_2
	Grupo control	(R)	O_3	—	O_4

Donde:

- R : asignación aleatoria de las unidades objeto de estudio a los grupos.
- X : variable independiente (programa de capacitación) grupo experimental.
- O_1 y O_2 : medición antes (O_1) y medición después (O_2) de la variable dependiente (volumen de las exportaciones) del grupo experimental.
- O_3 y O_4 : medición antes (O_3) y medición después (O_4) de la variable dependiente (volumen de exportaciones) del grupo control.

Análisis del diseño. Igual que el diseño anterior, en el que la conformación de los grupos se hace de manera aleatoria, los dos grupos pueden considerarse idénticos y, por tanto, las variables extrañas pueden afectarlos de manera similar. La diferencia de este diseño con el anterior radica en la medición previa que se le hace al grupo que recibe la capacitación. Entonces, podría afirmarse que la diferencia positiva a favor del grupo experimental respecto al grupo de control obedece a efectos del curso recibido.

Diseño de Solomon para cuatro grupos

Este diseño presenta las siguientes características:

Es un diseño que usa dos grupos experimentales y dos grupos control, con el propósito de controlar al máximo los efectos de las variables extrañas.

www.FreeLibros.me

El esquema de este diseño es:

Grupo experimental 1:	(*R*)	O_1	X	O_2
Grupo control 1:	(*R*)	O_3	-	O_4
Grupo experimental 2:	(*R*)	-	X	O_5
Grupo control 2:	(*R*)	-	-	O_6

El segundo grupo experimental no recibe prueba previa; no obstante, en lo demás es idéntico al primero. El segundo grupo control se somete a una sola medición de la variable dependiente (volumen de exportaciones) después de la prueba (programa de capacitación).

Análisis del diseño. De acuerdo con McDaniel y Gates (1999, p. 264), este diseño permite diversas mediciones del efecto de la variable independiente (programa de capacitación). Éstas son: $(O_2 - O_1) - (O_4 - O_3)$, $(O_6 - O_5)$ y $(O_2 - O_4)$. Si estas medidas concuerdan, las inferencias obtenidas acerca del efecto de la variable independiente serían bastante válidas y confiables. Además, este diseño permite medir directamente los efectos de la interacción de la variable independiente y los efectos previos a la medición $[(O_2 - O_4) - (O_5 - O_6)]$.

Los ejemplos anteriores permiten ilustrar cómo puede desarrollarse un mismo tema de investigación con diferentes grados de control de variables extrañas y, por tanto, diferentes grados de validez y confiabilidad en los resultados de la investigación.

Los diseños aquí ejemplificados son los de uso más frecuente. Sin embargo, existen otros diseños que igualmente pueden utilizarse para el desarrollo de una investigación. Tales diseños son:

Diseño experimental de series cronológicas

Se utiliza cuando el investigador está interesado en analizar efectos en el *mediano* y *largo plazos*, porque tiene bases para suponer que la influencia de la variable independiente sobre la dependiente tarda en manifestarse; por ejemplo, programas de capacitación profesional, difusión de innovación, aplicación de nuevas teorías económicas y administrativas a la actividad económica y administrativa del país o de las empresas.

Características. Se realizan varias mediciones sobre la variable dependiente durante un período prolongado. Las mediciones pueden ser tantas como se quieran y sea posible aplicar.

Se trabaja con dos o más grupos objeto de estudio; los sujetos o las unidades de análisis se asignan aleatoriamente a cada grupo.

Diseños factoriales

Estos diseños manipulan *dos* o *más variables* independientes e incluyen dos o más niveles de presencia en cada una de las variables independientes. Se utilizan con más frecuencia en investigación del comportamiento; por ejemplo, en una empresa, analizar el impacto de un ajuste en el salario, simultáneamente con el desarrollo de un programa de capacitación sobre el nivel de productividad; o el grado de ausentismo y satisfacción de los trabajadores.

www.FreeLibros.me

El número de grupos que se forman en un diseño factorial es igual a todas las posibles combinaciones que surjan al cruzar los niveles de una variable independiente con los niveles de las otras variables. Así, un diseño 2 × 2 (dos variables independientes y dos variables dependientes) requiere cuatro grupos; un diseño 3 × 2 (tres variables independientes y dos variables dependientes) requiere seis grupos.

Los sujetos o las unidades de prueba en los diseños factoriales se asignan aleatoriamente.

RESUMEN

La definición del diseño experimental de investigación está determinada por el tipo de investigación que va a realizarse y por la hipótesis que va a probarse durante el desarrollo de la investigación. Se habla de diseños cuando se hace referencia a la investigación experimental, que consiste en demostrar que la modificación de una variable (variable independiente) ocasiona un cambio predecible en otra (variable dependiente). La investigación experimental se realiza mediante los llamados diseños, que son un conjunto de procedimientos con los cuales se manipulan una o más variables independientes y se mide su efecto sobre una o más variables dependientes.

En la investigación experimental existen diversos tipos de diseño, que se clasifican de diferentes formas. Sin embargo, la clasificación más usada es la de Campbell y Stanley (1996), quienes identifican tres categorías generales de diseños de investigación: preexperimentales, cuasi experimentales y experimentales verdaderos. Esta clasificación se fundamenta en tres características básicas de los diseños: 1) el grado de control que se ejerce sobre las variables objeto de estudio, 2) el grado de aleatoriedad con que se asignan los sujetos de la investigación a un grupo o a varios de ellos y 3) la presencia o ausencia de grupos de control.

Los diseños preexperimentales presentan el más bajo control de variables y en éstos no se realiza asignación aleatoria de los sujetos al experimento; los diseños experimentales puros o verdaderos se caracterizan por un alto grado de control de las variables y porque existe asignación aleatoria de los sujetos a los grupos (experimental y control) participantes en la investigación.

Un aspecto importante en los diseños experimentales es su grado de validez, el cual se afecta de forma permanente y por consiguiente hay que estar atentos a ello. Al respecto, en los diseños se presentan dos clases fundamentales de validez, a las que hay que atender: la interna y la externa.

La validez interna mide si la acción de las variables independientes o los tratamientos producen los efectos en la variable dependiente. Si el investigador demuestra que la variable experimental (variable independiente) o de tratamiento produjo las diferencias observadas en la variable dependiente, se dice que el experimento tiene

www.FreeLibros.me

validez interna. La validez externa se centra en la posibilidad de que los resultados del experimento se generalicen a personas, medios y tiempos en el medio real.

Para el estudio de los experimentos se utiliza un sistema estándar de notación:

- *X* : indica la variable independiente (acción realizada sobre la población objeto del estudio); también se le conoce como tratamiento.
- *O* : se emplea para indicar la medición de la variable dependiente.
- *R* : se utiliza para indicar la asignación aleatoria de las unidades de prueba a los grupos en el experimento.
- *G* : indica el grupo objeto del estudio.
- *GE* : grupo experimental.
- *GC* : grupo control.

EJERCICIOS DE REPASO Y ANÁLISIS

1. ¿Qué se entiende por diseño de investigación experimental?
2. ¿Cómo se clasifican los diseños de investigación experimental y en qué consiste cada uno?
3. Describa la relación entre objetivos de la investigación, tipo de estudio, hipótesis y diseño.
4. Explique en qué consiste la validez interna y externa de los diseños de investigación experimental.
5. Describa cada una de las amenazas de la validez interna y externa de los diseños de investigación experimental.
6. ¿Qué es y cómo se hace el control de variables extrañas en los experimentos?
7. Haga una revisión bibliográfica en la biblioteca y revise material en el cual se muestre la aplicación de diseños de investigación experimental en el campo de su disciplina; después, en clase, comparta el material revisado con sus colegas.
8. Mediante consulta de revistas especializadas en el campo de su disciplina, indague por investigaciones recientes en las que se hayan utilizado diseños de investigación experimental y exponga los resultados de su indagación en el salón de clase.
9. Indague en su escuela o universidad sobre la existencia de investigaciones realizadas en años anteriores, en las cuales se hayan utilizado diseños de investigación experimental. Comparta la información obtenida con sus colegas de la clase.

www.FreeLibros.me

7.9 La población y la muestra objeto de estudio. Determinación

Una vez concebida la idea de investigación, y luego de tener claridad sobre el problema que se va a investigar, plantear los objetivos que se espera lograr, contar con una justificación para desarrollar el estudio, tener un fundamento teórico, plantear la hipótesis o las preguntas de investigación, definir el tipo y el diseño de la investigación, el otro aspecto para tener en cuenta es definir la *población* o *muestra* con la cual se desarrollará la investigación de interés.

En esta parte de la investigación, el interés consiste en definir quiénes y qué características deberán tener los sujetos (personas, organizaciones o situaciones y factores) objeto de estudio.

En seguida, sólo se plantean algunos conceptos muy relevantes para tener en cuenta en esta etapa del proceso de investigación (la persona interesada en profundizar en cada tema en particular puede consultar expertos y/o material especializado).

Nota: a los aspectos relacionados con la población y muestra, la recopilación de la información, el procesamiento de ésta, así como al análisis y la discusión de resultados, usualmente se les conoce como *estrategias metodológicas* de la investigación.

7.9.1 Población

De acuerdo con Fracica (1988), población es "el conjunto de todos los elementos a los cuales se refiere la investigación. Se puede definir también como el conjunto de todas las unidades de muestreo" (p. 36).

Según Jany (1994), población es "la totalidad de elementos o individuos que tienen ciertas características similares y sobre las cuales se desea hacer inferencia" (p. 48); o bien, unidad de análisis.

Las dos anteriores definiciones son igualmente válidas para el propósito del presente libro.

Por ello, para estos autores una definición adecuada de población debe realizarse a partir de los siguientes términos: *elementos*, *unidades de muestreo*, *alcance* y *tiempo*.

Si desea hacerse un análisis del sector del cuero y el calzado en su país, la población sería:

- Alcance: cinco principales ciudades capitales.
- Tiempo: de 1996 a 2004.
- Elementos: todas las empresas del sector del cuero y el calzado ubicadas en el territorio nacional.
- Unidades de muestreo: todas las empresas del sector del cuero y el calzado en el país.

www.FreeLibros.me

7.9.2 Marco muestral

Se refiere a la lista, el mapa o la fuente de donde pueden extractarse todas las unidades de muestreo o unidades de análisis en la población, y de donde se tomarán los sujetos objeto de estudio.

7.9.3 Muestra

Es la parte de la población que se selecciona, de la cual realmente se obtiene la información para el desarrollo del estudio y sobre la cual se efectuarán la medición y la observación de las variables objeto de estudio.

Pasos en la selección de una muestra Siguiendo el esquema de Kinnear y Taylor (1993), los siguientes son los pasos para definir una muestra:

1. Definir la población.
2. Identificar el marco muestral.
3. Determinar el tamaño de la muestra.
4. Elegir un procedimiento de muestreo.
5. Seleccionar la muestra.

Variables de la población y su medición Según Fracica (1988), "uno de los aspectos fundamentales para la realización de una investigación es la necesidad de conocer ciertas *características* de la población objeto de estudio", a las cuales "se les conoce como variables y pueden ser de tipo cuantitativo o cualitativo" (p. 46).

Estas variables se analizan a partir de sus necesidades, ya sea en términos de datos de promedios o totales para las variables cuantitativas, y de proporciones o totales para las variables cualitativas.

EJEMPLO 7.25 Variables cualitativas y cuantitativas de la población

Un estudio busca conocer la opinión de los usuarios respecto a un nuevo producto. La variable opinión es una variable cualitativa y los datos se analizarán en términos de proporciones. Por ejemplo, 20% de los encuestados opinarán favorablemente sobre la calidad del producto.

Un estudio interesado en conocer el volumen de producción del sector automotor y sus variaciones durante los últimos tres años.

La variable volumen de producción es una variable cuantitativa y su análisis se realizará en términos de promedios o de totales; por ejemplo, el volumen de producción del sector automotor de los últimos tres años fue de 30 000 vehículos, promedio anual; y la producción total ascendió a 90 000 vehículos durante los tres años.

www.FreeLibros.me

7.9.4 Tamaño de la muestra

En la investigación científica, el *t*amaño de la muestra debe estimarse siguiendo los criterios que ofrece la estadística, y por ello es necesario conocer algunas técnicas o métodos de muestreo.

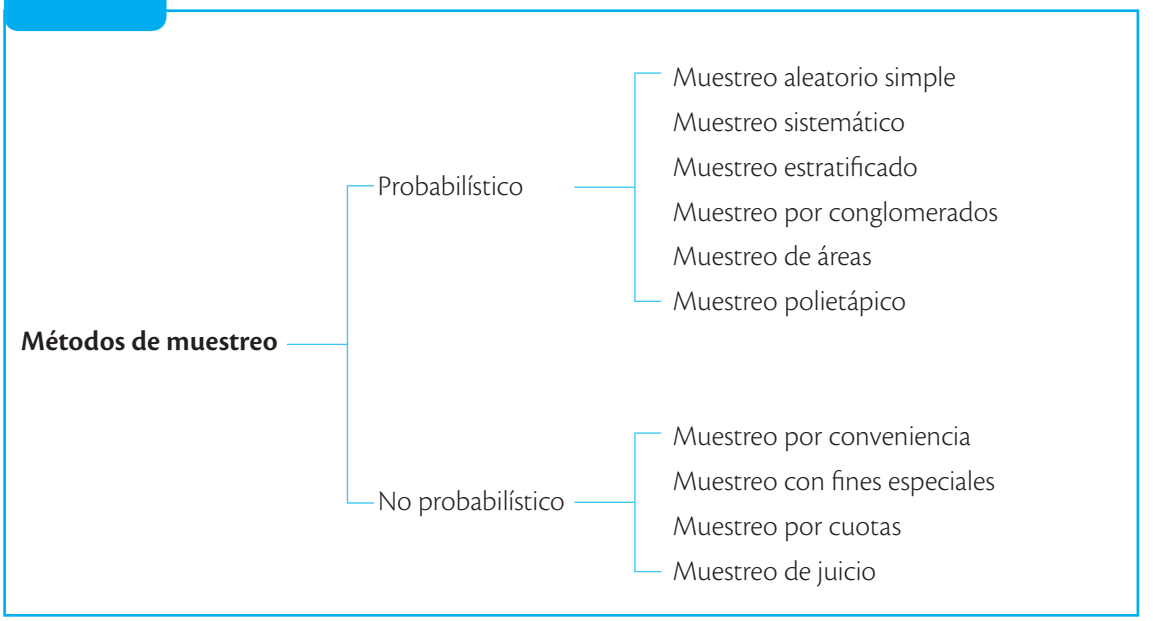
El método de muestreo utilizado para estimar el tamaño de una muestra depende del tipo de investigación que desea realizarse y, por tanto, de las hipótesis y del diseño de investigación que se hayan definido para desarrollar el estudio.

7.9.5 Métodos de muestreo

Existen varias clasificaciones para los métodos de muestreo. Según Weiers (1986), las más usadas son: diseños probabilísticos y no probabilísticos, y diseños por atributos y por variables. El primero de éstos es el más usual.

La Figura 7.13 sintetiza los métodos de muestreo más utilizados en la investigación científica para estimar tamaños de muestra en una población objeto de estudio.

FIGURA 7.13 Métodos de muestreo



Fuente: Weiers, R. (1986). *Investigación de mercados*. México: Prentice Hall.

www.FreeLibros.me

Así, de acuerdo con cada método de muestreo, existen criterios diferentes para estimar el tamaño de la muestra.

La Figura 7.14 ilustra la relación entre la población y muestra con los otros componentes de proceso de investigación científica según la espiral del proceso de investigación científica.

FIGURA 7.14 **El proceso de investigación. Determinación de la población y la muestra objeto de estudio**

www.FreeLibros.me

Casos y talleres

Cómo estimar el tamaño de una muestra estadística representativa de una población

Muestreo aleatorio simple (MAS)

En investigación, el muestreo aleatorio simple se utiliza cuando en el conjunto de una población, cualquiera de los sujetos tiene la variable o variables objeto de la medición.

Procedimiento para estimar el tamaño de muestra

Para estimar el tamaño de muestra representativo de una población, mediante el muestreo aleatorio simple (MAS), se sigue este procedimiento:

1. Se identifica si se trata de una población infinita (no se conoce el número total de la población *N*) o finita (se conoce el número total de la población *N*) y también si la variable objeto del estudio es cuantitativa o cualitativa. Este paso es importante para definir la fórmula por utilizar (la fórmula para estimar la muestra en una población infinita es diferente de la que se utiliza para estimar la muestra en poblaciones finitas) y la forma de estimar el valor del error de estimación que se va a utilizar en la respectiva fórmula (cuando la variable es cualitativa se utiliza un valor en porcentaje no superior a 10%, mientras que en el caso de las variables cuantitativas, el valor del error de estimación también es máximo 10%, pero es un valor absoluto estimado con base en el valor de la media).
2. Se determina la desviación estándar de la población que se puede obtener mediante resultados de estudio previos o mediante la realización de una "muestra piloto". Cuando se estima la desviación estándar de la población mediante la muestra piloto, se comienza por construir una tabla de frecuencias, resultado de la medición de la variable objeto del estudio (X_i).
3. Se define el nivel de confianza *Z* que se debe utilizar, el cual se encuentra en la *función de distribución normal tipificada* o *tabla Z*, que hoy día aparece en los libros de estadística y en varios de los libros de metodología de investigación. En este libro se encuentra como Anexo C. Para efectos de investigación, los valores de *Z* que se van a utilizar corresponden a niveles de confianza iguales o superiores a 90%, que en valores tipificados *Z* son valores iguales o superiores a 1,96.
4. Se define el error de estimación *E* (diferencia máxima entre la media muestral y la media poblacional que está dispuesto a aceptar el investigador, dependiendo del nivel de confianza (valores *Z*) y del valor estimado de la desviación *S*). En investigación científica, al error de estimación *E* se le asignan valores menores o iguales a 10% en relación inversa con los valores de *Z* y *S*, es decir, cuando se asigna un valor de nivel de confianza pequeño y la desviación estándar es grande, el error de estimación

www.FreeLibros.me

E* debe ser pequeño. El valor estimado del coeficiente de $S/\overline{X}$ permite definir cuándo la desviación estándar de una población es grande o pequeña. En los casos en los que el valor de $S/\overline{X}$ son menores o iguales que 30%, se dice que la desviación estándar es pequeña (se entiende que la respectiva población tiende a ser homogénea), mientras que cuando la relación $S/\overline{X}$ es mayor que 30% pero menor que 60%, se dice que la desviación estándar es grande (se considera que la respectiva población tiende a ser heterogénea), pero si la relación $S/\overline{X}$ es superior a 60%, entonces se dice que la población objeto del estudio es heterogénea y en ese caso se recomienda una muestra piloto mayor y asignar nivel de confianza superior a 95% y error de estimación pequeño –menor o igual que 3%.

5. Se calcula el valor del tamaño de muestra representativo según la fórmula correspondiente y se concluye sobre el tamaño mínimo de muestra requerido especificando los respectivos valores de nivel de confianza (Z), desviación estándar, estimado (S) y error de estimación utilizado (E).

Caso 7.1 Introducción de una nueva tarjeta de crédito- Banco JTRQP

Usted(es) ha(n) sido contratado(a)s por una reconocida entidad financiera (Banco JTRQP) con casa matriz en Estados Unidos, que quiere iniciar operaciones de banca en Colombia con motivo de la firma del TLC. El respectivo banco quiere realizar un estudio de mercados que le permita evaluar el gasto promedio mensual (en $) que actualmente realizan las personas (hombres y mujeres con edades entre 25 y 40 años) por utilización de tarjeta de crédito. Su función es apoyar al gerente de Marketing para estimar el tamaño de muestra representativo con el fin de realizar el respectivo estudio que tendrá lugar en la ciudad de *BGTX*. Por favor, estime(n) el tamaño de la respectiva muestra representativa para el grupo de población objetivo por encuestar.

Procedimiento

1. Se identifica si la población es infinita o finita (para el caso se estimará la muestra en las dos condiciones) y se define la variable objeto de medición (X_i) (en este caso, esa variable se denomina "Gasto promedio mensual por uso de tarjeta de crédito"), la cual es una variable cuantitativa y se define la fórmula del muestreo que se va a utilizar.
2. Se determina el valor de la desviación estándar, para el caso, valor que se obtendrá mediante muestra piloto: se mide la variable objeto del estudio en por lo menos 30 personas con las características de la respectiva población sobre la cual se realizará el estudio. En este caso, para estimar la desviación estándar se ha decidido realizar una muestra piloto a 57 personas con edades entre 25 y 40 años, residentes en la ciudad de *BGTX* y usuarios de tarjeta de crédito. A continuación, se muestra la tabla de frecuencias que se obtuvo de esa muestra piloto.

---

* *El valor del error de estimación E se define de forma diferente para las variables cuantitativas que para las variables cualitativas. En el caso de las variables cuantitativas, el valor de E se obtiene multiplicando el valor del promedio de la respectiva variable por el porcentaje definido para realizar el estudio en función de los valores del nivel de confianza Z y del valor de la desviación estándar estimado S, mientras que para las variables cualitativas, el valor de E se obtiene del valor en porcentaje que se defina, también en función de los valores de Z y S.*

www.FreeLibros.me

TABLA 7.4 **Gasto promedio mensual por uso de tarjeta de crédito en personas con edades entre 25 y 40 años en la ciudad de *BGTX***

No. de encuestas	x_i = gasto promedio/ mes (miles de $)	No. de encuestas	x_i = gasto promedio/ mes (miles de $)
1	200	30	1 300
2	450	31	900
3	370	32	1 200
4	1 000	33	350
5	800	34	250
6	400	35	1 100
7	300	36	700
8	250	37	500
9	900	38	700
10	1 100	39	800
11	750	40	500
12	500	41	300
13	300	42	450
14	500	43	800
15	800	44	700
16	750	45	300
17	350	46	250
18	1 200	47	300
19	250	48	750
20	600	49	900
21	750	50	600
22	900	51	1 200
23	450	52	900
24	1 100	53	1 100
25	650	54	1 250
26	350	55	900
27	500	56	700
28	900	57	300
29	450	Total: $n_0 = 57$	$\bar{X} = \sum x_i / n_0 = 663{,}51$

www.FreeLibros.me

3. Se estima el valor de la respectiva desviación estándar mediante la siguiente fórmula:

$$S = \sqrt{\sum_{i=1}^{n} (X_i - \bar{X})^2 / n_0 - 1} \quad ==> S = \$312{,}12$$

Donde:

S : desviación estándar de la muestra piloto de personas con edades entre 25 y 40 años, residentes en la ciudad de *BGTX* y usuarios de tarjeta de crédito.

X_i : gasto promedio de cada una de las personas encuestadas en la muestra piloto.

$\bar{X}$: gasto promedio del conjunto de las personas encuestadas de la muestra piloto.

n_0 : número de personas entrevistadas en la muestra piloto.

$$\text{Homogeneidad o heterogeneidad de la muestra} = \frac{S}{\bar{X}} ==> \frac{312{,}12}{663{,}51} > 30\%$$

En este caso, debido a que el valor de la relación $S/\bar{X}$ es superior al 30%, se considera que la población tiende a ser heterogénea y el valor que se asigne a E (error de estimación) se sugiere que debe ser pequeño, en lo posible menor que 5% (el valor máximo de E es 10%).

4. Estimación del tamaño de la muestra

a. Población infinita

En este caso, se supone que **no** se conoce el número total de personas (N) con edades entre 25 y 40 años, residentes en la ciudad de *BGTX* y usuarios de tarjeta de crédito; entonces, para estimar el tamaño de la muestra se utiliza la fórmula de muestreo aleatorio simple para población infinita, como se muestra a continuación:

$$n = \frac{Z^2 \times S^2}{E^2}$$

Donde:

n : tamaño necesario de la muestra de personas con edades entre 25 y 40 años, residentes en la ciudad de *BGTX* y usuarios de tarjeta de crédito.

Z : nivel de confianza o margen de confiabilidad (para este caso: 96% de confianza, Z = 2,054).

S : desviación estándar de la población (estimada mediante una muestra piloto aplicada a 57 personas con edades entre 25 y 40 años, residentes en la ciudad de *BGTX* y usuarios de tarjeta de crédito) ==> S = 312,12.

www.FreeLibros.me

E : error de estimación (en este caso, se tomó un error de estimación pequeño debido a que la población objeto del estudio tiende a ser heterogénea en cuanto a gasto a través del uso de tarjeta de crédito. Entonces, $E = 0{,}03$ del promedio del gasto en el respectivo uso de tarjeta).

$$n = \frac{Z^2 \times S^2}{E^2}, \text{ ==> } n = \frac{(2{,}054)^2 \times (312{,}12)^2}{(0{,}03 \times 663{,}51)^2} \text{ ==> } n = 1\,038 \text{ personas}$$

5. **En conclusión**, utilizando un nivel de confianza del 96% y un error de estimación del 3% se requiere entrevistar a por lo menos 1 038 personas de entre 25 y 40 años, residentes en la ciudad de *BGTX*, usuarios de tarjeta de crédito, para conocer el gasto promedio mensual en tarjeta de crédito de las personas en el respectivo rango de edad.

b. Tamaño de muestra cuando la población es finita (se conoce *N*)

En el caso que se conozca el total de la población, es decir, el número total de personas con edades entre 25 y 40 años, residentes en la ciudad de *BGTX* y usuarios de tarjeta de crédito, para estimar el tamaño de la muestra se utiliza la fórmula de muestreo aleatorio simple, como se muestra a continuación:

$$n = \frac{S^2}{\frac{E^2}{Z^2} + \frac{S^2}{N}}$$

Donde:

N : número total de personas entre 25 y 40 años de edad, residentes en la ciudad de *BGTX* y usuarios de tarjeta de crédito. En este caso se supone un número total de 780 000 personas con el respectivo perfil.

n : tamaño necesario de la muestra de personas de entre 25 y 40 años de edad, residentes en la ciudad de *BGTX* y usuarios de tarjeta de crédito.

Z : nivel de confianza o margen de confiabilidad (en este caso: 96% de confianza, $Z = 2{,}054$).

S : desviación estándar de la población (estimada mediante una muestra piloto aplicada a 57 personas de entre 25 y 40 años de edad, residentes en la ciudad de *BGTX* y usuarios de tarjeta de crédito) ==> $S = 312{,}12$.

E : error de estimación (en este caso, se tomó un error de estimación pequeño debido a que la población objeto del estudio tiende a ser heterogénea en cuanto a gasto a través del uso de tarjeta de crédito: $E = 0{,}03$* del promedio del gasto en el respectivo uso de tarjeta).

* *Recuérdese que para el caso de estimación de tamaños de muestra donde la variable objeto de medición es una variable cuantitativa, el error de estimación se obtiene mediante el porcentaje (máximo 10%) de la media de la muestra piloto de la población.*

www.FreeLibros.me

El procedimiento para la estimación del respectivo tamaño de muestra es el mismo realizado para estimar tamaño de muestra para la población infinita, pero la fórmula es diferente y como sigue:

$$n = \frac{S^2}{\frac{E^2}{Z^2} + \frac{S^2}{N}} \quad n = \frac{(312{,}12)^2}{\frac{(0{,}03 \times 663{,}51)^2}{(2{,}054)^2} + \frac{(312{,}12)^2}{780\,000}} \quad n = 1\,036 \text{ personas}$$

En conclusión, utilizando un nivel de confianza del 96% y un error de estimación del 3% para una población de 780 000 personas con edades de entre 25 y 40 años, residentes en la ciudad de *BGTX*, usuarios de tarjeta de crédito, se requiere entrevistar a por lo menos 1 036 personas para conocer el gasto promedio mensual en tarjeta de crédito por parte de las personas en el respectivo rango de edad en la mencionada ciudad.

Caso 7.2 Percepción sobre la gestión del conocimiento

Suponga que usted tiene interés en realizar un estudio cuyo objetivo es conocer la percepción que tienen las personas que laboran en una reconocida compañía del sector farmacéutico en esta ciudad, sobre la forma como en ella se gestiona el conocimiento. Su función es estimar el tamaño de muestra representativa de trabajadores (directivos y no directivos) de la respectiva compañía a quienes para tal propósito se les aplicará una encuesta conformada por 58 enunciados estructurados en escala Likert, donde el encuestado califica cada enunciado según su percepción, en puntuaciones de 1 a 5, siendo:

1 = Pésima; 2 = Mala; 3 = Regular; 4= Buena; 5 = Excelente.

Procedimiento

1. Se identifica si la población es infinita o finita. Para el caso, se puede saber el número total de personas que trabajan en la compañía en la ciudad, y por ello se dirá que es una población finita. La variable objeto de medición (X_i) –en este caso se denomina "Percepción sobre la forma de gestionar el conocimiento"– es una variable cualitativa* y es un criterio clave para definir el error de estimación**.
2. Se determina el valor de la desviación estándar, para el caso, valor que se obtendrá mediante muestra piloto: se mide la variable objeto del estudio en por lo menos 30 personas con las características de la respectiva población sobre la cual se realizará el estudio. En este caso, para estimar la desviación estándar se aplica la encuesta a una muestra piloto de 40 personas que laboren en la compañía. El siguiente es el respectivo cuadro de frecuencias de la muestra piloto sobre la percepción de las personas encuestadas sobre la forma en que se gestiona el conocimiento en la compañía. En este caso se utilizó la distribución unidimensional, pero el CD que acompaña a este libro ilustra un ejemplo más completo.

* *En el CD que acompaña a este libro se ejemplifica la estimación de una muestra representativa.*

** *Recuérdese que para el caso de estimación de tamaños de muestra donde la variable objeto de medición es una variable cualitativa, el error de estimación se define como un valor porcentaje (máximo 10%) para variables cualitativas estructuradas en escala Likert.*

www.FreeLibros.me

TABLA 7.5 **Percepción sobre gestión del conocimiento. Compañía farmacéutica**

No. de encuestas	X_i Percepción	No. de encuestas	X_i Percepción
1	3	24	2
2	2	25	1
3	5	26	1
4	3	27	2
5	5	28	1
6	1	29	4
7	2	30	4
8	5	31	4
9	4	32	3
10	3	33	2
11	2	34	2
12	2	35	5
13	2	36	4
14	1	37	4
15	3	38	4
16	2	39	3
17	5	40	3
18	4	41	4
19	4	42	4
20	3	43	1
21	2	44	2
22	2	45	2
23	5	Total n_0 = 45	$\bar{X} = \sum x_i / n_0$ = 2,92

La fórmula para la estimación de la desviación estándar es la siguiente:

$$S = \sqrt{\sum_{i=1}^{n} (X_i - \bar{X})^2 / n_0 - 1} \quad \text{==> } S = \$1{,}29$$

Donde:

S : desviación estándar de la percepción sobre la forma como se gestiona el conocimiento en la compañía por las personas de la muestra piloto.

X_i : percepción de cada una de las personas encuestadas en la muestra piloto sobre la forma como se gestiona el conocimiento en la compañía.

www.FreeLibros.me

$\overline{X}$: percepción promedio del conjunto de las personas encuestadas de la muestra piloto sobre la forma como se gestiona el conocimiento en la compañía.

n_0 : número de personas entrevistadas en la muestra piloto.

$$\text{Homogeneidad o heterogeneidad de la muestra} = \frac{S}{\overline{X}} \Longrightarrow \frac{1{,}29}{2{,}92} > 30\%$$

En este caso, debido a que el valor de la relación $S/\overline{X}$ es superior al 30%, se considera que la población tiende a ser heterogénea y el valor que se asigne a E (error de estimación) se sugiere que debe ser pequeño, en lo posible menor que 5% (el valor máximo de E es 10%).

3. Se estima el tamaño representativo de la muestra. Se sabe que la compañía tiene una nómina de 7 584 trabajadores, es decir, este es el número total de personas que laboran en la compañía; por tanto, es una población finita y la fórmula de muestreo aleatorio simple, que se debe utilizar para estimar el tamaño de muestra representativa de personas que laboran en la compañía por encuestar para conocer la percepción sobre la forma como se gestiona el conocimiento, se presenta a continuación:

$$n = \frac{S^2}{\frac{E^2}{Z^2} + \frac{S^2}{N}}$$

Donde:

N : número total de personas que laboran en la compañía (7 584).

n : tamaño necesario de la muestra representativa de personas que se debe encuestar para determinar la verdadera percepción que tienen las personas en la compañía sobre la forma como se gestiona el conocimiento en la misma.

Z : nivel de confianza o margen de confiabilidad (para este caso a las directivas de la compañía les interesa ser rigurosas y, por ello, deciden utilizar un nivel de confianza alto: 95% de confianza, para un valor de $Z = 1{,}96$).

S : desviación estándar de la población (estimada mediante la muestra piloto aplicada a 45 personas ==> $S = 1{,}29$)

E : error de estimación (en este caso se tomó un error de estimación un valor pequeño, debido a que la población objeto del estudio tiende a ser heterogénea en cuanto a la percepción de la variable objeto de medición; entonces: $E = 0{,}05$*****.

$$n = \frac{S^2}{\frac{E^2}{Z^2} + \frac{S^2}{N}} \quad n = \frac{(1{,}29)^2}{\frac{(0{,}05)^2}{(1{,}96)^2} + \frac{(1{,}29)^2}{7\,584}} \quad n = 1\,912 \text{ personas}$$

* *Recuérdese que para el caso de estimación de tamaños de muestra donde la variable objeto de medición es una variable cualitativa, el error de estimación se define como un valor porcentaje (máximo 10%).*

www.FreeLibros.me

En conclusión, utilizando un nivel de confianza del 95% y un error de estimación del 5% para el total de trabajadores de la compañía (7 584 personas), se requiere encuestar a por lo menos 1 912 personas para conocer la percepción representativa de los trabajadores sobre la forma como se gestiona el conocimiento en la compañía.

Caso 7.3 Percepción sobre calidad y pertinencia de las lecturas complementarias en un determinado programa académico

Con el propósito de complementar el ejemplo anterior que está desarrollado con una distribución unidimensional (una sola pregunta para todas las personas), éste es más completo y se desarrolla aprovechando el aplicativo Process en el CD que acompaña a este libro y que también ilustra casos tanto del muestreo aleatorio simple, como del muestreo proporcional y del muestreo estratificado.

El director del programa de una determinada facultad desea realizar una investigación cuyo objetivo es evaluar la actitud de los estudiantes de la facultad o programa respecto de la calidad y la pertinencia de las lecturas complementarias. Estas las recomiendan los docentes de las asignaturas o núcleos temáticos correspondientes a los distintos semestres que conforman el programa. El instrumento (cuestionario) que se va a utilizar en la recolección de la información es una encuesta conformada por 50 enunciados en escala Likert, validada tanto por un comité de jueces como por una prueba piloto, con las categorías de respuesta siguientes:

Totalmente de acuerdo	5
Parcialmente de acuerdo	4
Indiferente	3
Parcialmente en desacuerdo	2
Totalmente en desacuerdo	1

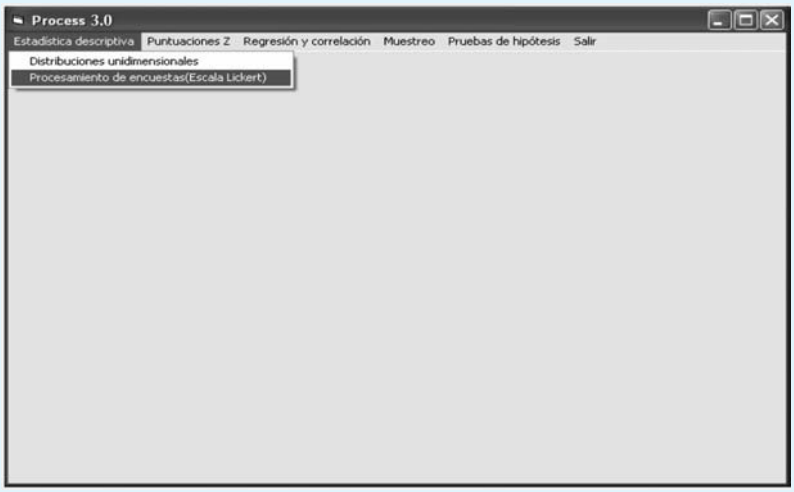
TALLER 7.1 Procesamiento de la información apoyado en el programa Process

Para desarrollar estos casos, usted puede apoyarse en la herramienta para el procesamiento de información Process, que se incluye en el CD. A continuación, se describen las instrucciones para definir las encuestas en escala de Likert.

1. Haga clic en Process 3.0, ubicado en la sección Herramientas y ejercicios del CD.
2. En la barra de menú haga clic en **Estadística descriptiva** y luego haga clic en **Procesamiento de encuestas (Escala Likert)**. (En la pantalla 7.1 se muestra el menú de la aplicación).

www.FreeLibros.me

Pantalla 7.1 Barra de menú Process



3. En la pantalla 7.2 se presentan las condiciones generales para procesar este tipo de encuestas. De acuerdo con cada una de las opciones, ingrese la información que se relaciona a continuación y cuando haya terminado haga clic en el botón **OK**.

   Proyecto nuevo

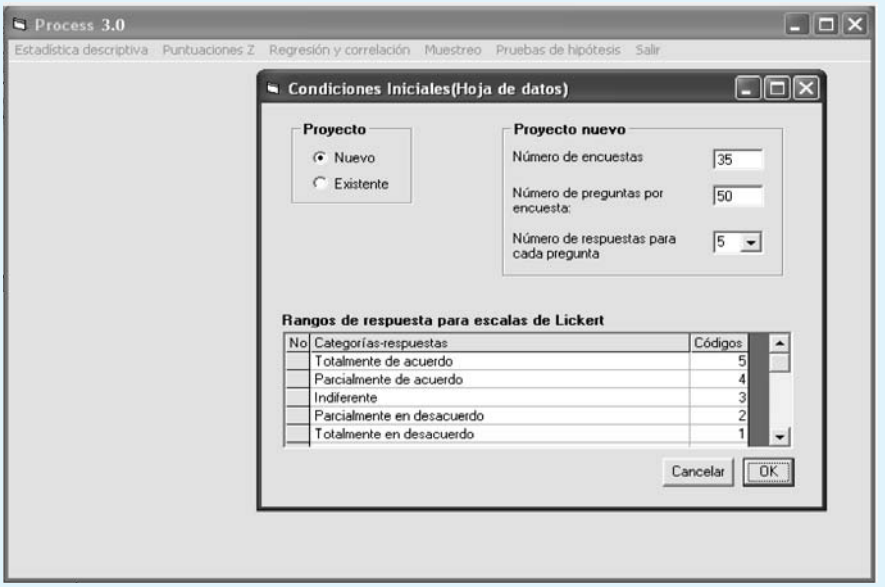
Número de encuestas	35
Número de preguntas por encuesta	50
Número de respuestas para cada pregunta	5

4. Dependiendo de la selección que realice en el cuadro **Número de respuestas para cada pregunta**, el cuadro de diálogo **Rango de respuestas para escalas de Likert** se actualiza, proponiendo una serie de rangos de respuesta y la codificación numérica que utilizará el programa para realizar los cálculos (columna **Códigos**). Si en el cuadro **Número de respuestas para cada pregunta** selecciona las opciones 2, 3 ó 5, el programa ajusta automáticamente el cuadro de diálogo **Rangos de respuesta para escalas de Likert**; en los demás casos digite el texto correspondiente a cada una de las categorías de respuesta. Si las opciones de respuesta que usted requiere en su análisis no concuerdan con las opciones propuestas, usted podrá modificar el texto de las opciones de respuesta propuestas. Para ello, haga clic sobre la categoría que desea modificar y digite el nuevo texto.

www.FreeLibros.me

Pantalla 7.2 Condiciones iniciales (Hoja de datos)

Si va a ingresar nuevos datos, en el cuadro **Proyecto** *seleccione la opción* **Nuevo**. *Si ha guardado previamente sus datos en un archivo, seleccione la opción* **Existente** *y ubique el nombre del archivo.*



El cuadro **Número de preguntas por encuesta** *se relaciona directamente con el cuadro de diálogo* **Rangos de respuestas para escalas de Likert**. *Si no aparecen opciones de respuesta, digítelas manualmente.*

El programa propone la codificación que se muestra en la columna **Códigos** *para tabular las respuestas. Trabaje preferiblemente con éstas para evitar resultados inesperados.*

Procedimientos para el cálculo del tamaño de la muestra

1. Se plantea la fórmula estadística apropiada para estimar el tamaño de muestra

$$n = \frac{S^2}{\frac{E^2}{Z^2} + \frac{S^2}{N}} \text{ (Tamaño de población conocido)}$$

$$n = \frac{S^2 \times Z^2}{E^2} \text{ (Cuando no se conoce el total de la población)}$$

Donde:

n : tamaño de la muestra por estimar.
S : desviación estándar de la población.
Z : margen de confianza.
E : error de estimación o diferencia máxima entre la media muestral y la media poblacional que el investigador está dispuesto a aceptar con el nivel de confianza que se defina.
N : número total de la población objeto del estudio.

www.FreeLibros.me

2. Se estima el valor de S (desviación estándar): el valor de S se puede obtener de dos formas: (a) conocida o estimada a partir de estudios anteriores y (b) estimación mediante una muestra piloto.

   En este caso no existen estudios previos, por lo que es necesario realizar una *muestra piloto* para estimar el valor de S; este valor se estima de dos formas:(a) estimando la desviación estándar para cada una de la preguntas del cuestionario para todos los encuestados y tomar el mayor valor de las diferentes desviaciones por pregunta para estimar el tamaño de la muestra, o (b) estimando la desviación estándar del promedio de total de las encuestas de todos los encuestados, como se ilustra a continuación, mediante la siguiente fórmula:

$$S = \sqrt{\sum_{i=1}^{n} (X_i - \bar{X})^2 / n_0 - 1}$$

   Donde:

   S : desviación estándar de la población a partir de la muestra piloto.

   X_i : puntuación o medición de la variable objeto de estudio (en este ejemplo, medición de la actitud respecto a la lectura complementaria) para cada uno de los estudiantes de la muestra piloto.

   $\bar{X}$: promedio de la población o de la medición de la variable objeto de estudio de cada uno de los estudiantes de la muestra piloto.

   n_o : número de estudiantes de la muestra piloto (los criterios estadísticos indican que la muestra piloto de un estudio la constituyen un número igual o superior a 30 sujetos tomados aleatoriamente de la población y sobre los cuales se recoge la información para estimar el valor de S).

   Veamos entonces cómo se estima el valor de S para este caso mediante muestra piloto.

   Se aplica la encuesta de las 50 preguntas a un número igual o superior a 30 estudiantes (en este caso se aplicará a 35 estudiantes) seleccionados aleatoriamente y luego se calcula la puntuación promedio de cada una de las encuestas aplicadas a los 35 estudiantes y ese valor es el valor de los X_i. (Para realizar estos cálculos hay que apoyarse en la herramienta de procesamiento de información Process, incluida en el CD).

TALLER 7.2 **Procesamiento de encuestas**

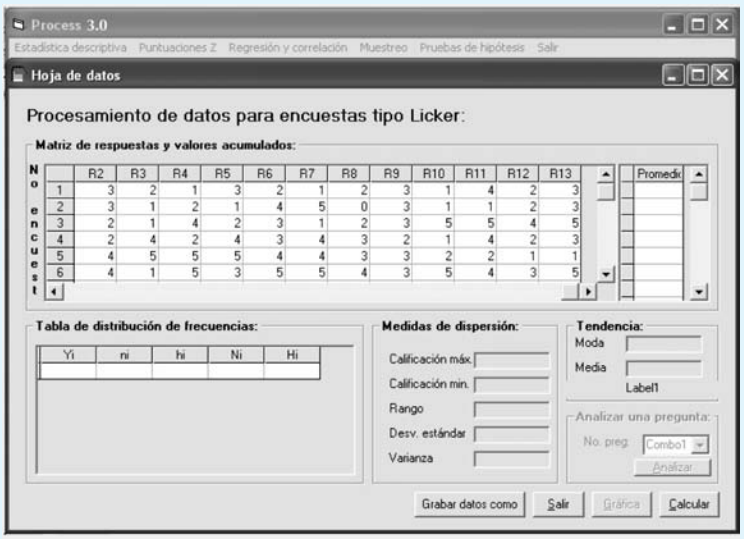
1. En la **Matriz de respuestas y valores acumulados**, digite los resultados correspondientes a las preguntas de cada una de las encuestas así: la encuesta 1 se diligencia en dirección izquierda-derecha, en el primer renglón, digitando los códigos de las respuestas definidos en la pantalla 7.2. Por ejemplo, si usted seleccionó la opción 2 del cuadro **Número de respuestas para cada pregunta**, las **Categorías- respuesta** son: **Sí**, **No**, **Pregunta no contestada**. Por tanto, los códigos que se deben digitar en la matriz son valores entre 0 y 2, pues así se codificó.

www.FreeLibros.me

Los títulos que aparecen en la parte superior (en el encabezado de cada columna) representan el número de la pregunta; por ejemplo, 1 representa la respuesta codificada a la pregunta R1 (un valor entre 0 y 2); 2 representa el código de respuesta a la pregunta R2, y así sucesivamente, hasta finalizar la totalidad de la encuesta. En la pantalla 7.3 se ilustra la matriz diligenciada.

2. Desplace el cursor de una casilla a otra con las teclas de dirección (flechas derecha-izquierda y arriba-abajo).
3. Para realizar los cálculos, una vez se ingresen las respuestas de todas las encuestas en las casillas de la matriz, haga clic en el botón **Calcular.**

Pantalla 7.3 Matriz de respuesta y valores acumulados



Digite aquí las respuestas codificadas, de acuerdo con los códigos de la pantalla 7. 2.

Con la opción **Grabar datos como**, *usted puede archivar sus datos y recuperarlos posteriormente.*

Una vez que haya digitado la totalidad de los datos, haga clic en el botón **Calcular**.

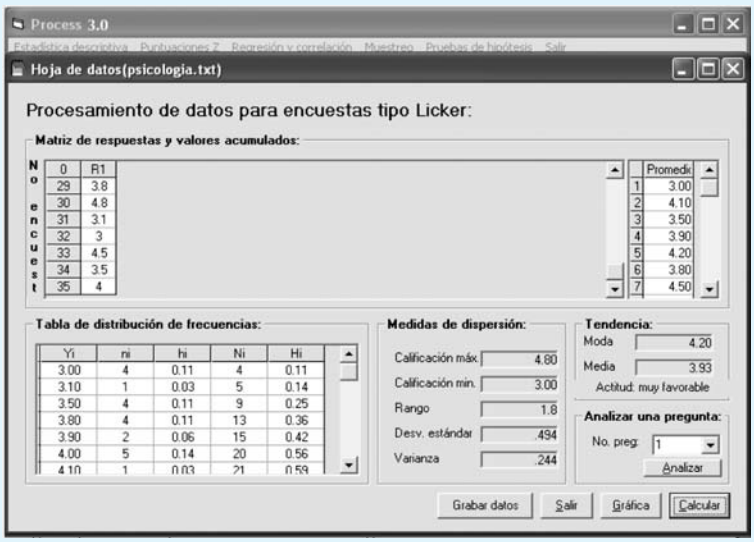
Muy bien, ya se ha procesado parte de la información. La pantalla 7.4 contiene el valor correspondiente a la desviación estándar, la misma que se necesitará posteriormente para calcular el tamaño definitivo de la muestra.

Las operaciones necesarias para determinar el promedio de calificación de cada estudiante ya han sido realizadas por el programa, así como la aplicación de las fórmulas para definir el valor de la media y de la moda.

Desde este punto, usted también puede examinar el comportamiento de una única pregunta. Seleccione la pregunta que desea analizar y haga clic sobre el botón **Analizar**.

www.FreeLibros.me

Pantalla 7.4 Comportamiento de datos agrupados, desviación estándar y análisis de preguntas



Para analizar el comportamiento de una única pregunta, seleccione el número correspondiente a la pregunta que desea y haga clic sobre el botón **Analizar**.

En la tabla de distribución de frecuencias puede observar el comportamiento de los datos agrupados.

El valor de la desviación estándar servirá posteriormente para calcular el tamaño de la muestra.

Procedimientos para el cálculo del tamaño de la muestra (continuación)

La desviación estándar indica el grado de dispersión de la muestra en relación con la media; es decir, el grado de homogeneidad o heterogeneidad de la muestra. Cuando la S = grande, la población tiende a ser heterogénea y cuando la S = pequeña, la población tiende a ser homogénea.

3. Se define el nivel de confianza en valores Z. El nivel de confianza o valor de Z es cualquier valor de la tabla para las puntuaciones Z, que son valores normalizados. Sin embargo, en investigación estadísticamente se suele tomar valores para niveles de confianza igual o superior al 90%. Para este estudio el investigador ha seleccionado un nivel de confianza del 95%, por lo que Z = 1,96.
4. Se define el valor del error de estimación E. El valor de E se define a partir de la desviación estándar (S) estimada a partir de la muestra piloto, y usualmente se recomienda darle al error de estimación un valor inferior a 10% (cuando la variable objeto de medición es cualitativa sólo se coloca el valor del porcentaje definido, pero cuando la respectiva variable es cuantitativa el valor de E es igual al porcentaje asignado multiplicado por el valor de la media (promedio) de la muestra).

www.FreeLibros.me

Si la desviación estándar es grande en proporción con la media, el error de estimación (E) debe ser pequeño, tendiendo a cero (0); si el valor de la desviación estándar (S) es pequeño en relación con la media, entonces el valor de error de estimación (E) puede ser cercano pero no superior a 10% del valor de la media o promedio.

En esta investigación, como el valor de la desviación estándar es pequeño e indica que la población de estudiantes tiende a ser homogénea, entonces el valor del error de estimación puede ser $E = 0{,}06$ ó $E = 6\%$ ó $0{,}06$.

5. Se estima el tamaño de la muestra.

$$n = \frac{S^2}{\frac{E^2}{Z^2} + \frac{S^2}{N}} \qquad n = \frac{(0{,}49)^2}{\frac{(0{,}06)^2}{(1{,}96)^2} + \frac{(0{,}49)^2}{N}}$$

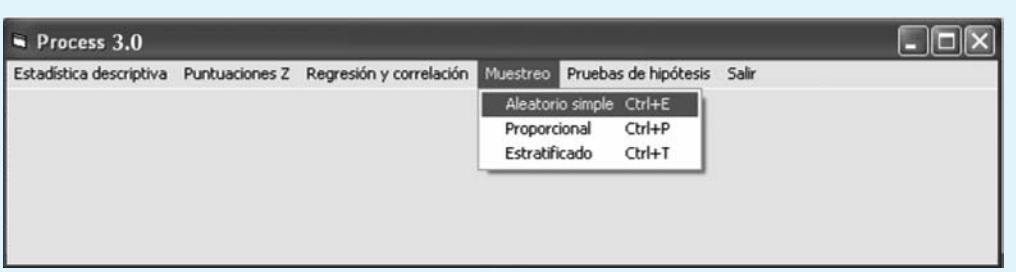
El valor de N (total de estudiantes) se obtuvo de los archivos que reposan en la Secretaría Académica de la Facultad para el total de alumnos matriculados en el programa objeto de estudio del año en curso. ($N = 1\,200$).

En conclusión, para realizar el estudio de evaluación de la actitud de los estudiantes del programa de psicología, respecto a la calidad y pertinencia de las lecturas complementarias recomendadas por los profesores de las diferentes áreas del conocimiento de la carrera, se requiere una muestra representativa de 214 estudiantes, con un nivel de confianza del 95% y error de estimación del 0,06 ó 6%.

TALLER 7.3 Determinación del tamaño de la muestra (MAS)

Para determinar el tamaño de la muestra con el método de muestreo aleatorio simple, abra la aplicación Process y en la barra de menú haga clic en **Muestreo**, luego haga clic en la opción **Aleatorio simple**.

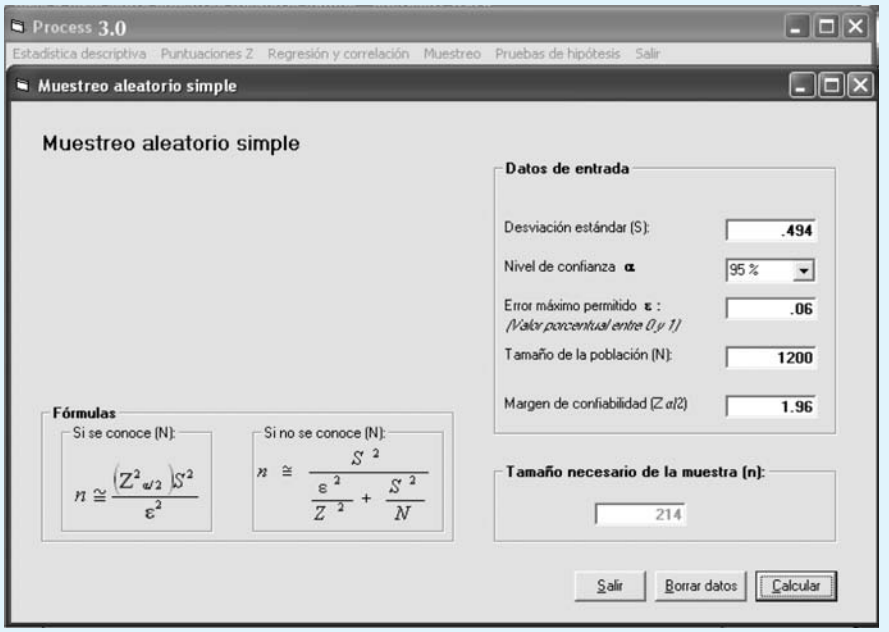
Pantalla 7.5 Opciones de muestreo



Ahora que ya definió los valores necesarios para realizar estos cálculos, digítelos en cada una de las casillas correspondientes. Una vez complete los datos, haga clic sobre el botón **Calcular.**

www.FreeLibros.me

Pantalla 7.6 Margen de confiabilidad



El valor correspondiente al margen de confiabilidad se ajusta automáticamente, de acuerdo con el nivel de confianza dado.

Muestreo proporcional

El muestreo proporcional se utiliza cuando la variable objeto de la medición se mide en proporciones o probabilidades de ocurrencia.

Procedimiento para estimar el tamaño de muestra representativo

Para estimar el tamaño de muestra representativa de una población, con muestreo proporcional, igual que en el muestreo aleatorio simple, se procede de la siguiente manera:

1. Se identifica si se trata de una población infinita (no se conoce el número total de la población *N*) o finita (cuando se conoce el número total de la población *N*). Recuérdese que este paso es importante para definir la fórmula que se debe utilizar, ya que es distinta la fórmula de tamaño de muestra para poblaciones infinitas que para finitas.
2. Se determina la proporción de la población que tiene la variable objeto de medición, la cual se puede obtener mediante resultados de estudios previos o mediante la realización de una "muestra piloto". Cuando se estima la proporción de la población mediante la muestra piloto, se comienza por construir una tabla de frecuencias, resultado de la medición de la variable objeto del estudio (P_i).

www.FreeLibros.me

3. Se define el nivel de confianza *Z* por utilizar, el cual se encuentra en la *función de distribución normal tipificada* o *Tabla Z*, igual como se hace para el muestreo aleatorio simple. Para efectos de investigación, los valores de *Z* por utilizar corresponden a niveles de confianza iguales o superiores a 90% que en valores tipificados *Z* son valores iguales o superiores a 1,64.
4. Se define el valor del error de estimación *E* (diferencia máxima entre la proporción muestral y la proporción poblacional que el equipo investigador está dispuesto aceptar en función del nivel de confianza definido para el estudio). En muestreo proporcional, el valor de *E* debe ser menor o igual que 6% ó 0,06, en relación directa con los valores de *Z*, es decir, cuando se define un valor de nivel de confianza bajo cercano al 90% el error de estimación *E* debe ser pequeño.
5. Se calcula el valor del tamaño de muestra según la fórmula correspondiente a la población (infinita o finita) y se concluye sobre el tamaño mínimo de muestra representativo especificando los valores asignados al nivel de confianza (*Z*), la proporción estimada (*P*) y el error de estimación (*E*) utilizado.

Para ejemplificar la estimación de un tamaño de muestra representativo con muestreo proporcional, a continuación se presenta el caso 7.4, tanto para población infinita como para población finita y simulando cuando no se conoce el valor de la proporción (*P*) y se estima por muestra piloto, cuando el valor de *P* se obtiene de estudios se toma dicho valor y cuando no se conoce *P*, se puede asignarle a la respectiva proporción el valor de 0,5 ó 50%.

Caso 7.4 Mujeres cabeza de familia cotizantes a salud

Una organización encargada de la protección social en la ciudad está interesada en estimar el tamaño de muestra representativo que le permita conocer la proporción de personas mujeres cabeza de familia de estrato 1 y 2, que son cotizantes para los servicios de salud de los miembros de su familia.

Procedimiento para estimar el tamaño de la muestra representativa para una población infinita

El siguiente es el procedimiento para la estimación de la respectiva muestra cuando la población es infinita y no se tiene referencia de una proporción de mujeres cabeza de familia cotizantes de servicios de salud para los miembros de su familia en su ciudad.

1. Se construye la tabla de distribución de frecuencias para las mujeres cabeza de familia de los estratos 1 y 2, que son cotizantes para los servicios de salud para los miembros de su familia en la ciudad. Para ello, se realizará una muestra piloto de mujeres de la población objeto del estudio para saber la proporción de mujeres cabeza de familia que cotizan servicios de salud. Después se les pregunta a las mujeres seleccionadas si cotizan (1) o no (2) como cabeza de familia.

www.FreeLibros.me

TABLA 7.6 **Proporción de mujeres cabeza de familia en los estratos 1 y 2 que cotizan servicio de salud para los miembros de su familia**

No. de encuestas	Cotiza salud	No. de encuestas	Cotiza salud
1	1	32	2
2	1	33	1
3	2	36	2
4	1	35	1
5	2	36	1
6	1	37	1
7	1	38	1
8	1	39	1
9	1	40	1
10	2	41	2
11	2	42	2
12	1	43	2
13	1	44	1
14	1	45	1
15	2	46	1
16	2	47	1
17	1	48	1
18	2	49	1
19	1	50	1
20	1	51	1
21	1	52	2
22	1	53	1
23	1	54	1
24	1	55	1
25	2	56	2
26	2	57	2
27	2	58	2
28	2	59	1
29	2	60	1
30	2	61	1
31	2	62	1

www.FreeLibros.me

No. de encuestas	Cotiza salud	No. de encuestas	Cotiza salud
63	2	70	1
64	2	71	1
65	1	72	1
66	1	73	2
67	2	74	2
68	2	75	1
69	2		
		Total No. = 75	$P = \frac{\text{Total enc., con 1}}{\text{Total de encuestas}}$ $P = 50/75 = 0{,}666$

2. Se recurre a la fórmula para estimar el tamaño de muestra.

$$n = \frac{Z^2 \times P \times Q}{E^2}$$

Donde:

n : tamaño de muestra por estimar (número de mujeres cabeza de familia de estrato 1 y 2 que es necesario encuestar para conocer la proporción representativa de mujeres del estrato 1 y 2 que cotizan a salud como cabeza de familia).

Z : nivel de confianza o margen de confiabilidad (97%, es decir, $Z = 2{,}17$).

P : proporción de mujeres del estrato 1 y 2 de la muestra piloto que cotizan a salud como cabeza de familia.

$Q = 1 - P$: proporción de mujeres del estrato 1 y 2 de la muestra piloto que no cotizan a salud como cabeza de familia.

E : error de estimación (diferencia máxima entre la proporción muestral y la proporción poblacional que el equipo investigador está dispuesto a aceptar en función del nivel de confianza definido para el estudio). En este caso, $E = 0{,}03$ ó 4%. Entonces 3%.

$$n = \frac{Z^2 \times P \times Q}{E^2} \quad n = \frac{2{,}17^2 \times 0{,}666 \times 0{,}334}{0{,}03^2} \quad n = 1\,164$$

En conclusión, con un nivel de confianza del 97%, un error de estimación del 3% es necesario encuestar un total de 1 164 mujeres de estrato 1 y 2 cotizantes a salud para conocer la proporción representativa de mujeres de los respectivos estratos, en la ciudad que cotizan a salud como cabeza de familia.

www.FreeLibros.me

Procedimiento para estimar el tamaño de la muestra representativa para una población finita

El procedimiento para estimar el tamaño de muestra representativo es igual que para las poblaciones infinitas, pero la fórmula es diferente, así:

$$n = \frac{Z^2 \times P \times Q \times N}{E^2(N\text{-}1) + Z^2 \times P \times Q}$$

Donde:

- n : tamaño de muestra por estimar (número de mujeres cabeza de familia de estrato 1 y 2 que es necesario encuestar para conocer la proporción representativa de mujeres del estrato 1 y 2 que cotizan a salud como cabeza de familia).
- Z : nivel de confianza o margen de confiabilidad (97%, es decir, $Z = 2{,}17$).
- P : proporción de mujeres del estrato 1 y 2 de la muestra piloto que cotizan a salud como cabeza de familia.
- $Q = 1\text{-}P$: proporción de mujeres del estrato 1 y 2 de la muestra piloto que no cotizan a salud como cabeza de familia.
- N : número total de mujeres del estrato 1 y 2 que cotizan a salud en la ciudad.
- E : error de estimación (diferencia máxima entre la proporción muestral y la proporción proporcional que el equipo investigador está dispuesto aceptar en función del nivel de confianza definido para el estudio). En este caso $E = 0{,}03$ ó 3%.

$$n = \frac{Z^2 \times P \times Q \times N}{E^2(N\text{-}1) + Z^2 \times P \times Q} \qquad n = \frac{2{,}17^2 \times 0{,}666 \times 0{,}334 \times 580\,000}{0{,}03^2\,(580\,000 - 1) + 2{,}17^2 \times 0{,}666 \times 0{,}334}$$

$n = 1\,162$ mujeres de los estratos 1 y 2 que cotizan a salud en la respectiva ciudad.

En conclusión, para una población de 580 000 mujeres de los estratos 1 y 2, con un nivel de confianza del 97% y un error de estimación del 3%, es necesario encuestar un total de 1 162 mujeres de estrato 1 y 2 cotizantes a salud para conocer la proporción representativa de mujeres de los respectivos estratos en la ciudad que cotizan a salud como cabeza de familia.

Muestreo estratificado

En investigación científica se utiliza el muestreo estratificado para estimar tamaños de muestra cuando en el estudio la población objeto de la investigación existe alguna variable altamente correlacionada con la variable objeto de medición y esa variable hace que los sujetos sean altamente heterogéneos, por lo que se recomienda agrupar los respectivos sujetos por rangos o estratos. Por ejemplo, alguien quiere conocer la capacidad de compra de un determinado producto y sabe que la variable correlacionada con la capacidad de consumo es el ingreso y éste es distinto para las personas de esa población; entonces, en ese caso, es necesario estratificar los sujetos de la población por rangos de ingreso.

www.FreeLibros.me

En otro caso, se requiere conocer la percepción sobre la prestación de un servicio de diversión que incluye adultos y niños; entonces, la edad influye en la percepción y por ello se recomienda agrupar la población por rangos de edad.

El procedimiento para estimar el tamaño de muestra en el muestreo estratificado es el siguiente:

1. Se identifica si existe alguna variable correlacionada con la variable objeto de medición que haga que los sujetos de la población sean heterogéneos respecto a esa variable.
2. Se determina si la población es infinita o finita (para el caso se estimará la muestra únicamente para el caso de una población finita) y se define la variable objeto de medición (Xi).
3. Se determina el tamaño total de la muestra (el cual se estima de la misma forma que para el muestreo aleatorio simple) estimando la desviación estándar de la población mediante una muestra piloto; asignando el respectivo nivel de confianza Z que se va a utilizar (niveles de confianza iguales o superiores a 90% que en valores tipificados Z son valores iguales o superiores a 1,64 y el error de estimación E), valores menores o iguales a 10% en relación inversa con los valores de Z y S.
4. Se estratifica la población según la variable correlacionada con la variable objeto de la medición y que hace que los sujetos sean heterogéneos entre sí.
5. Se estima el valor del tamaño de la muestra representativo para cada estrato a partir del tamaño de muestra total de la población según la fórmula correspondiente del muestreo estratificado, especificando los respectivos valores de nivel de confianza (Z), desviación estándar estimada (S) para cada estrato.

Caso 7.5 Inversión en libros de texto por estudiantes universitarios en la ciudad *XTG*, año 2010

La empresa editorial RGTBF S. A. está interesada en realizar una investigación sobre la adquisición de libros de texto por estudiantes universitarios de los programas de psicología en la ciudad *XTG*, en el año 2010. Para ello ha decidido estimar un tamaño de muestra de estudiantes vinculados a los programas de psicología (se sabe que los programas son ofrecidos por universidades con valor de matrícula claramente diferenciada), por lo cual considera conveniente estratificar a los estudiantes por estratos de matrículas, así: "universidades tipo A" aquellas con valor de matrícula más alto, "universidades tipo B" las de valor de matrícula medio y "universidades tipo C" las de valor de matrícula más bajo.

www.FreeLibros.me

Estimación del tamaño de muestra para el total de la población a partir de una muestra piloto

TABLA 7.7 **Inversión en libros de texto por parte de estudiantes universitarios en la ciudad *XTG*, año 2010.**

No. de encuestas	Inversión en libro/ semestre (miles de $)	No. de encuestas	Inversión en libro/ semestre (miles de $)
1	1 500	33	850
2	200	34	600
3	750	35	600
4	200	36	600
5	600	37	1 100
6	350	38	450
7	120	39	600
8	1 300	40	1 200
9	1 700	41	950
10	250	42	1 300
11	180	43	120
12	450	44	540
13	590	45	610
14	800	46	900
15	900	47	750
16	1 170	48	225
17	200	49	690
18	150	50	350
19	450	51	900
20	600	52	810
21	400	53	450
22	350	54	345
23	270	55	210
24	620	56	190
25	1 600	57	1 100
26	1 330	58	900
27	1 250	59	700
28	245	60	100

www.FreeLibros.me

No. de encuestas	Inversión en libro/ semestre (miles de $)	No. de encuestas	Inversión en libro/ semestre (miles de $)
29	460	61	210
30	700	62	350
31	1500	63	500
32	900	Total No. = 63	$\overline{X} = \Sigma x_i/No.$ $\overline{X} = 655,3$

El valor de la desviación estándar estimado a partir de una muestra piloto es:

$$S = \sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2 / n_0 - 1} \Longrightarrow S = \$\,412,87$$

Para la estimación del total de la muestra piloto de la población se sabe que el total de los estudiantes matriculados en los respectivos programas académicos de las universidades objeto del estudio para el año 2010 es de 9 700 en la mencionada ciudad; entonces, la fórmula para estimar el total de la muestra es:

$$n = \frac{S^2}{\frac{E^2}{Z^2} + \frac{S^2}{N}}$$

Donde:

N : número total de estudiantes matriculados para el año 2010 en los programas de psicología en la ciudad en las universidades objeto del estudio (9 700).

n : tamaño total necesario de la muestra de estudiantes matriculados en los programas de psicología que adquieren libros de texto por período académico para sus estudios.

Z : nivel de confianza o margen de confiabilidad (en este caso: 96% de confianza, $Z = 2,054$).

S : desviación estándar de la población (estimada mediante una muestra piloto, $S = \$412,87$).

E : error de estimación (en este caso, se tomó un error de estimación pequeño, debido a que la población objeto del estudio tiende a se heterogénea en cuanto a inversión para adquisición de libros; entonces $E = (0,03 \times 655,32)$*.

Homogeneidad o heterogeneidad de la muestra $= \frac{S}{\overline{X}} \Longrightarrow \frac{412,87}{655,32} > 30\%$

* *Recuérdese que para el caso de estimación de tamaños de muestra donde la variable objeto de medición es una variable cuantitativa, el error de estimación se obtiene mediante el porcentaje (máximo 10%) de la media de la muestra piloto de la población.*

www.FreeLibros.me

El procedimiento para la estimación del respectivo tamaño de muestra es el mismo realizado para estimar el tamaño de muestra para la población infinita, pero la fórmula, diferente, es la siguiente:

$$n = \frac{S^2}{\frac{E^2}{Z^2} + \frac{S^2}{N}} \quad n = \frac{(412{,}87)^2}{\frac{(0{,}03 \times 655{,}32)^2}{(2{,}054)^2} + \frac{(412{,}87)^2}{9\,700}} \quad n = 1\,561 \text{ personas}$$

La editorial requiere encuestar a por lo menos 1 561 estudiantes matriculados en los programas de psicología en la ciudad para el período académico de 2010. A partir de este tamaño de muestra, se debe estimar el tamaño de muestra por estrato para cada uno de los tres tipos de universidades, según valor pagado por matrícula.

Para estimar el tamaño de la muestra por tipo de universidad, se utiliza la siguiente fórmula para cada uno de los estratos o rangos en los que se agrupa la población objeto de la investigación (las fórmulas están presentadas de forma general para cualquier número de rangos o estratos)

$$n_A = \frac{(n)(N_A)(S_A)}{(N_A)(S_A) + (N_B)(S_B) + (N_C)(S_C) + (N_D)(S_D) + \ldots\ldots + (N_X)(S_X)}$$

$$n_B = \frac{(n)(N_B)(S_B)}{(N_A)(S_A) + (N_B)(S_B) + (N_C)(S_C) + (N_D)(S_D) + \ldots\ldots + (N_X)(S_X)}$$

$$n_C = \frac{(n)(N_C)(S_C)}{(N_A)(S_A) + (N_B)(S_B) + (N_C)(S_C) + (N_D)(S_D) + \ldots\ldots + (N_X)(S_X)}$$

$$n_D = \frac{(n)(N_D)(S_D)}{(N_A)(S_A) + (N_B)(S_B) + (N_C)(S_C) + (N_D)(S_D) + \ldots\ldots + (N_X)(S_X)}$$

$$n_X = \frac{(n)(N_X)(S_X)}{(N_A)(S_A) + (N_B)(S_B) + (N_C)(S_C) + (N_D)(S_D) + \ldots\ldots + (N_X)(S_X)}$$

Donde:

n : tamaño total de la muestra representativa (para este caso n = 1 561 estudiantes) estimada mediante muestra piloto con nivel de confianza del 97% y un error de estimación del 3%.

n_A : tamaño de la muestra para el estrato A (tamaño de muestra de estudiantes de psicología de las universidades tipo A).

n_B : tamaño de la muestra para el estrato B (tamaño de muestra de estudiantes de psicología de las universidades tipo B).

n_C : tamaño de la muestra para el estrato C (tamaño de muestra de estudiantes de psicología de las universidades tipo C).

www.FreeLibros.me

n_C : tamaño de la muestra para el estrato D (para el caso objeto del estudio no hay estrato D).

n_X : tamaño de la muestra para el estrato X (para el caso objeto del estudio no hay estrato X).

N_A : total de la población del estrato A (total de estudiantes matriculados en programas de psicología en las universidades tipo A para el año 2010, NA = 1 300 estudiantes matriculados).

S_A : desviación estándar de la población del estrato A (la desviación se estima a partir de muestra piloto para la población del respectivo estrato, para este caso la S_A = \$520).

N_B : total de la población del estrato B (total de estudiantes matriculados en programas de psicología en las universidades tipo B para el año 2010, N_A = 2 800 estudiantes matriculados)

S_B : desviación estándar de la población del estrato B (la desviación se estima a partir de muestra piloto para la población del respectivo estrato, en este caso la S_B = \$370)

N_C : total de la población del estrato C (total de estudiantes matriculados en programas de psicología en las universidades tipo C para el año 2010, N_C = 5 600 estudiantes matriculados).

S_C : desviación estándar de la población del estrato C (la desviación se estima a partir de muestra piloto para la población del respectivo estrato, en este caso la S_C = \$150).

N_D : total de la población del estrato D (en el caso objeto del estudio no hay estrato D).

S_D : desviación estándar de la población del estrato D (en el caso objeto del estudio no hay estrato D).

N_X : total de la población del estrato *X* (en el caso objeto del estudio no hay estrato X).

S_X : desviación estándar de la población del estrato *X* (en el caso objeto del estudio no hay estrato X).

TABLA 7.8 **Agrupación de estudiantes de los programas de psicología por tipo de universidad**

Tipo de universidad según valor de matrícula	No. de estudiantes matriculados/ tipo de universidad	Desviación estándar por tipo de universidad (miles de \$)
Universidad tipo A	1 300	\$520
Universidad tipo B	2 800	\$370
Universidad tipo C	5 600	\$150

Entonces:

$$n_A = \frac{(n)(N_A)(S_A)}{(N_A)(S_A) + (N_B)(S_B) + (N_C)(S_C)}$$

$$n_A = \frac{(1\,561)(1\,300)(520)}{(1\,300)(520) + (2\,800)(370) + (5\,600)(150)} = 413 \text{ estudiantes}$$

www.FreeLibros.me

$$n_B = \frac{(n)(N_B)(S_B)}{(N_A)(S_A) + (N_B)(S_B) + (N_C)(S_C)}$$

$$n_B = \frac{(1\,561)(2\,800)(370)}{(1\,300)(520) + (2\,800)(370) + (5\,600)(150)} = 634 \text{ estudiantes}$$

$$n_C = \frac{(n)(N_C)(S_C)}{(N_A)(S_A) + (N_B)(S_B) + (N_C)(S_C)}$$

$$n_C = \frac{(1\,561)(5\,600)(150)}{(1\,300)(520) + (2\,800)(370) + (5\,600)(150)} = 514 \text{ estudiantes}$$

En conclusión, a fin de conocer la capacidad de compra representativa de los estudiantes de los programas de psicología para adquisición de libros en cada uno de los tipos de universidad según el valor de la matrícula en la ciudad, la editorial con un nivel de confianza del 96% y un error de estimación del 3% debe encuestar por lo menos a 413 estudiantes matriculados en universidades con valor de matrícula considerada alta, 634 estudiantes matriculados en universidades con valor de matrícula considerada media y 514 estudiantes matriculados en universidades con valor de matrícula considerada baja.

RESUMEN

Una vez concebida la idea de investigación, y luego de tener claridad sobre el problema que se va a investigar, hay que plantear los objetivos que se espera lograr, contar con una justificación para desarrollar el estudio, tener un fundamento teórico, plantear la hipótesis o las preguntas de investigación, definir el tipo y el diseño de la investigación, el; otro aspecto para tener en cuenta es definir la *población* o *muestra* con la cual se desarrollará la investigación de interés.

Al respecto, población es la totalidad o el conjunto de todos los sujetos o elementos que tienen ciertas características similares y a los cuales se refiere la investigación. La muestra es la *parte* de la población que se selecciona, de la cual realmente se obtiene la información para el desarrollo del estudio y sobre la cual se efectuarán la medición y la observación de las variables objeto de estudio. Los pasos para definir una muestra son los siguientes:

- Definición de la población.
- Identificación del marco muestral.
- Determinación del tamaño de la muestra.
- Elección de un procedimiento de muestreo.
- Selección de la muestra.

www.FreeLibros.me

En la población o muestra es necesario conocer ciertas *características*, a las cuales se les denomina como variables y pueden ser de tipo cuantitativo o cualitativo. Estas variables se analizan a partir de sus necesidades, ya sea en términos de datos de promedios o totales para las variables cuantitativas, y de proporciones o totales para las variables cualitativas.

En la investigación científica, el *tamaño de la muestra* debe estimarse siguiendo los criterios que ofrece la estadística, y por ello hay que conocer algunas técnicas o métodos de muestreo. El método de muestreo utilizado para estimar el tamaño de una muestra depende del tipo de investigación que desea realizarse y de las hipótesis y del diseño de investigación que se hayan definido para desarrollar el estudio.

En este libro se ha ilustrado la estimación de tamaño de muestra a partir de tres tipos de muestreo: muestreo aleatorio simple (MAS), muestreo proporcional y muestreo estratificado. Para el muestreo aleatorio simple y para el proporcional se ilustró el cálculo del tamaño de la muestra para el caso de poblaciones finitas e infinitas.

EJERCICIOS DE REPASO Y ANÁLISIS

1. ¿Qué es la población en una investigación? ¿ Qué es la muestra en una investigación?
2. ¿Cómo se define el tamaño de la muestra representativa de la población objeto de estudio o investigación que se va a realizar?
3. Defina un tema de investigación, establezca el objetivo general y dé un ejemplo de la estimación o del cálculo del tamaño de muestra para desarrollar la investigación.
4. Suponga que va a hacer una investigación para conocer la opinión de los clientes respecto a la calidad del servicio prestado por el restaurante JPL, al cual asisten diariamente, en promedio, 370 personas. ¿De qué tamaño debe ser la muestra de clientes que se deben entrevistar para que la información obtenida sea representativa?
5. Suponga que va a realizar un diagnóstico sobre la competitividad de las empresas del sector de las artes gráficas. El número de empresas del sector es de 3 900 empresas, discriminadas así: 2 180 son pequeñas empresas, 980 son medianas empresas y 740 grandes empresas. ¿De qué tamaño debe ser la muestra de empresas que se va a diagnosticar por cada tamaño de empresas, si se decide tomar una muestra total de 490 empresas?
6. Revise algún material bibliográfico sobre el tema de muestreo aplicado al campo de su profesión, repase los diferentes tipos de muestreo y haga ejercicios para compartir con lo(a)s compañero(a)s de clases.
7. Plantee situaciones del campo de su disciplina o profesión donde necesite estimar el tamaño de muestra, luego calcule el mismo para las diferentes situaciones planteadas.

www.FreeLibros.me

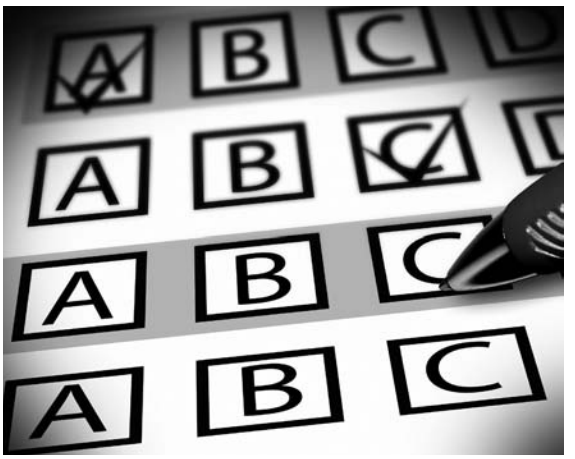
7.10 Obtención de la información. Recopilación

Un aspecto muy importante en el proceso de una investigación tiene relación con la obtención de la información, pues de ello dependen la confiabilidad y validez del estudio. Obtener información confiable y válida requiere cuidado y dedicación.

Esta etapa de recolección de información en investigación se conoce también como trabajo de campo.

Estos datos o información que va a recolectarse son el medio a través del cual se prueban las hipótesis, se responden las preguntas de investigación y se logran los objetivos del estudio originados del problema de investigación.

Los datos, entonces, deben ser confiables, es decir, deben ser pertinentes y suficientes, para lo cual es necesario definir las fuentes y técnicas adecuadas para su recolección.



Nota: en el caso del documento del anteproyecto, se debe mencionar cuáles serán las fuentes y las técnicas para la recolección de la información en el trabajo de campo, así como el proceso que se utilizará para tal efecto. En el anteproyecto se menciona qué se hará, mientras que en el documento del informe final se menciona lo que se hizo.

7.10.1 Fuentes de recolección de información

De acuerdo con Cerda (1998), usualmente se habla de dos tipos de *fuentes* de recolección de información: las *primarias* y las *secundarias*.

Fuentes primarias Son todas aquellas de las cuales se obtiene información directa, es decir, de donde se origina la información. Es también conocida como información de primera mano o desde el lugar de los hechos. Estas fuentes son las personas, las organizaciones, los acontecimientos, el ambiente natural, etcétera.

www.FreeLibros.me

Se obtiene información primaria cuando se observan directamente los hechos (presenciar una huelga, observar sistemáticamente el lugar de trabajo, etcétera), cuando se entrevista directamente a las personas que tienen relación directa con la situación objeto del estudio (en el caso de que quiera conocerse la opinión de los gerentes sobre el impacto de las medidas económicas en la actividad de las empresas, la información directa se genera cuando se entrevista directamente a los gerentes, y no cuando se lee en un periódico, un libro o se escucha en un noticiero).

Fuentes secundarias Son todas aquellas que ofrecen información sobre el tema que se va a investigar, pero que no son la fuente original de los hechos o las situaciones, sino que sólo los referencian. Las principales fuentes secundarias para la obtención de la información son los libros, las revistas, los documentos escritos (en general, todo medio impreso), los documentales, los noticieros y los medios de información.

En investigación, cualquiera de estas fuentes es válida siempre y cuando el investigador siga un procedimiento sistematizado y adecuado a las características del tema y a los objetivos, al marco teórico, a las hipótesis, al tipo de estudio y al diseño seleccionado.

En investigación, cuanto mayor rigor y exigencia se involucren en el proceso del desarrollo del estudio, más válido y confiable será el conocimiento generado.

7.10.2 Técnicas de recolección de información

En la actualidad, en investigación científica hay gran variedad de técnicas o instrumentos para la recolección de información en el trabajo de campo de una determinada investigación. De acuerdo con el método y el tipo de investigación que se va a realizar, se utilizan unas u otras técnicas.

Según Muñoz Giraldo et al. (2001), la investigación cuantitativa utiliza generalmente los siguientes instrumentos y técnicas para la recolección de información:

- Encuestas
- Entrevistas
- Observación sistemática
- Escalas de actitudes
- Análisis de contenido
- Test estandarizados y no estandarizados
- Grupos focales y grupos de discusión
- Pruebas de rendimiento
- Inventarios
- Fichas de cotejo
- Experimentos
- Técnicas proyectivas
- Pruebas estadísticas

www.FreeLibros.me

Según los mencionados autores, la investigación de tipo cualitativo utiliza sobre todo los siguientes instrumentos o técnicas, de acuerdo con el problema objeto de la investigación que se va a realizar:

- Entrevista estructurada y no estructurada
- Observación sistemática y no sistemática
- Historias de vida
- Autobiografías
- Anécdotas
- Relatos
- Notas de campo
- Preguntas etnográficas
- Análisis de documentos
- Diarios
- Cuadernos
- Archivos
- Cuestionarios
- Métodos sociométricos
- *Survey* social
- Inventarios y listados de interacciones
- Grabaciones en audio y video
- Fotografías y diapositivas
- Test de rendimiento
- Técnicas proyectivas
- Grupos focales y grupos de discusión

En concordancia con los autores citados, no todos los instrumentos o las técnicas se aplican a toda investigación. Sin embargo, la tendencia es utilizar baterías (aplicación de varios instrumentos que se complementen) a las diferentes investigaciones.

La tendencia contemporánea de la investigación es el diálogo entre enfoques, porque se acepta el criterio de que no hay métodos ni técnicas autosuficientes para la comprensión de la realidad de ningún aspecto u objeto de estudio; por tanto, se reconoce la necesidad de la complementariedad de métodos y técnicas.

A continuación se hace una presentación general de las principales técnicas o instrumentos de recolección de información en un proceso de investigación. Estas técnicas tienen aplicación en cualquiera de los enfoques cualitativo y cuantitativo de la investigación; por su importancia en la investigación actual, las mismas se presentan de manera más amplia en el capítulo 9 de este libro.

www.FreeLibros.me

Encuesta Es una de las técnicas de recolección de información más usadas, a pesar de que cada vez pierde mayor credibilidad por el sesgo de las personas encuestadas.

La encuesta se fundamenta en un cuestionario o conjunto de preguntas que se preparan con el propósito de obtener información de las personas.

Entrevista Técnica orientada a establecer contacto directo con las personas que se consideren fuente de información. A diferencia de la encuesta, que se ciñe a un cuestionario, la entrevista, si bien puede soportarse en un cuestionario muy flexible, tiene como propósito obtener información más espontánea y abierta. Durante la misma, puede profundizarse la información de interés para el estudio.

Observación directa Cada día cobra mayor credibilidad y su uso tiende a generalizarse, debido a que permite obtener información directa y confiable, siempre y cuando se haga mediante un procedimiento sistematizado y muy controlado, para lo cual hoy están utilizándose medios audiovisuales muy completos, especialmente en estudios del comportamiento de las personas en sus sitios de trabajo.

Análisis de documentos Técnica basada en fichas bibliográficas que tienen como propósito analizar material impreso. Se usa en la elaboración del marco teórico del estudio.

Para una investigación de calidad, se sugiere utilizar simultáneamente dos o más técnicas de recolección de información, con el propósito de contrastar y complementar los datos.

Internet No existe duda sobre las posibilidades que hoy ofrece Internet como una técnica de obtener información; es más, se ha convertido en uno de los principales medios para recabar información.

7.10.3 Proceso para la recolección de datos

La recopilación de información es un proceso que implica una serie de pasos. Aquí se presenta un esquema general que puede usarse para la recolección de los datos necesarios, para responder a los objetivos y para probar la hipótesis de la investigación, o ambos.

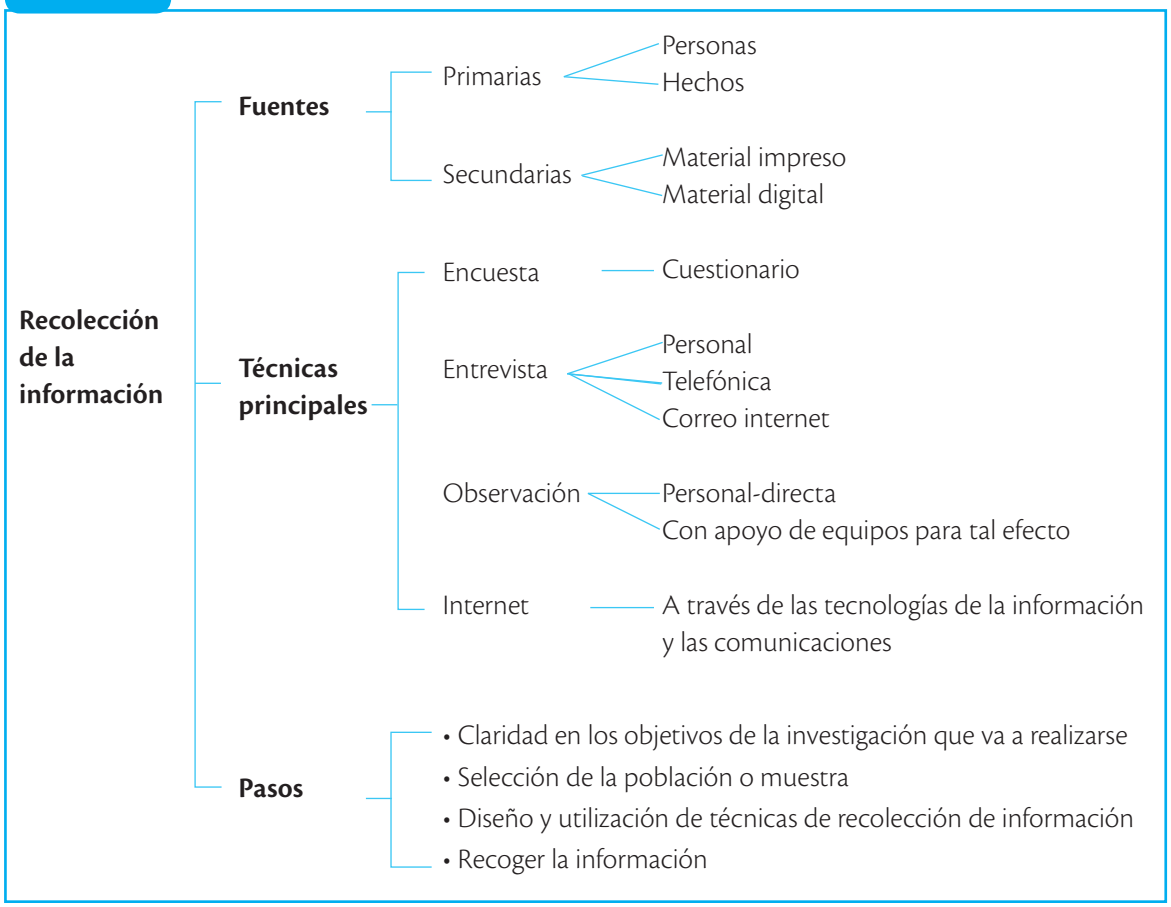
Estos pasos son los siguientes:

1. Tener claros los objetivos propuestos en la investigación y las variables de la hipótesis (si las hay).
2. Haber seleccionado la población o muestra objeto del estudio.
3. Definir las técnicas de recolección de información (elaborarlas y validarlas).
4. Recoger la información para luego procesarla para su respectiva descripción, análisis y discusión.

La tabla 7.9 sintetiza las principales fuentes y técnicas de obtención de información en el proceso de investigación científica.

www.FreeLibros.me

TABLA 7.9 **Fuentes y técnicas de obtención de información**



La Figura 7.15 ilustra la relación entre la recopilación de la información y otros aspectos del espiral del proceso de investigación científica.

www.FreeLibros.me

FIGURA 7.15 El proceso de investigación. Recopilación de la información

RESUMEN

En investigación es común referirse a dos tipos de fuentes de recolección de información: las *primarias* y las ***secundarias***. Las *fuentes primarias* son todas aquellas de las cuales se obtiene información directa, es decir, de donde se origina la información. Las *fuentes secundarias* son todas aquellas que ofrecen información sobre el tema que se va a investigar, pero que no son la fuente original de los hechos o las situaciones, sino los referenciales.

En investigación, cualquiera de estas fuentes es válida siempre y cuando el investigador siga un procedimiento sistematizado y adecuado a las características del tema y a los objetivos, al marco teórico, a las hipótesis, al tipo de estudio y al diseño seleccionado. Cuanto mayor rigor y exigencia se involucren en el proceso de recolección de la información para una investigación, más válido y confiable será el conocimiento generado.

En cuanto a técnicas de recolección de información, en la actualidad, en investigación científica hay gran variedad de técnicas o instrumentos para la recolección de información en el trabajo de campo de una determinada investigación. De acuerdo con el método y el tipo de investigación que se va a realizar, se utilizan unas u otras técnicas.

www.FreeLibros.me

Sin embargo, la tendencia es utilizar baterías (aplicación de varios instrumentos que se complementen) a las diferentes investigaciones.

Los pasos que deben seguirse en el proceso de recolección de información se pueden sintetizar en los siguientes:

- Clarificación de los objetivos propuestos en la investigación y las variables de la hipótesis (si las hay).
- Selección de la población o muestra objeto del estudio.
- Definición de las técnicas de recolección de información (elaborarlas y validarlas).
- Recolección de la información para luego procesarla para su respectiva descripción, análisis y discusión.

EJERCICIOS DE REPASO Y ANÁLISIS

1. ¿Qué significa recopilar información en un proceso de investigación científica?
2. ¿Qué requisitos debe tener la información para que sea considerada por el investigador o tenida en cuenta por el investigador?
3. ¿Qué se entiende por fuentes primarias y secundarias en un proceso de obtención de información para una investigación científica?
4. ¿Qué se entiende por técnicas de obtención de información en un proceso de investigación científica y cuáles son las principales técnicas que se utilizan para tal efecto?
5. ¿Cuáles son las principales técnicas o los instrumentos de recolección de información en los enfoques cuantitativo y cuantitativo de la investigación científica?
6. ¿Cuál es la tendencia actual en el uso de las técnicas o los instrumentos de recolección de información de la investigación científica?
7. Consulte material bibliográfico sobre fuentes, técnicas y procedimientos para recoger información en un proceso de investigación científica y discuta los resultados de esta averiguación con sus colegas o con su profesor(a) de investigación.
8. Seleccione un tema de investigación del campo de su profesión, defina los objetivos que tendría el estudio y describa cuáles serían las fuentes, las técnicas y el procedimiento que se seguiría para realizar la respectiva investigación.
9. Consulte revistas especializadas en el campo de su profesión y revise informes de investigación allí publicados donde se describan las fuentes, las técnicas y el procedimiento utilizado para obtener la información para la respectiva investigación. Comparta los resultados de su consulta con los compañeros de clases.

www.FreeLibros.me

7.11 Procesamiento de la información. Datos

Esta parte del proceso de investigación consiste en procesar los datos (dispersos, desordenados, individuales) obtenidos de la población objeto de estudio durante el trabajo de campo, y tiene como finalidad generar resultados (datos agrupados y ordenados), a partir de los cuales se realizará el análisis según los objetivos y las hipótesis o preguntas de la investigación realizada, o de ambos.

El procesamiento de datos debe realizarse mediante el uso de herramientas estadísticas con el apoyo del computador, utilizando alguno de los programas estadísticos que hoy fácilmente se encuentran en el mercado.

Nota: en el documento del *anteproyecto* se debe mencionar cómo se tiene previsto efectuar el procesamiento de la información obtenida en el trabajo de campo, mientras que en el *informe final* hay que mostrar este procedimiento mediante la descripción de los resultados de la investigación.

7.11.1 Pasos para el procesamiento de datos

Para efectuar un procesamiento de datos se deben seguir estos pasos:

1. Obtener la información de la población o muestra objeto de la investigación.
2. Definir las variables o los criterios para ordenar los datos obtenidos del trabajo de campo.
3. Definir las herramientas estadísticas y el programa de cómputo que va a utilizarse en el procesamiento de datos.
4. Introducir los datos en el computador y activar el programa para que procese la información.
5. Imprimir los resultados.

7.11.2 Herramientas estadísticas para el procesamiento de resultados

El procesamiento de resultados puede efectuarse mediante*:

Análisis de Pareto Técnica para estudiar fuentes de problemas y las prioridades relativas de sus causas. Se emplea frecuentemente para evaluar causas de problemas de calidad en programas de *total quality management* (TQM).

Diagrama de causa/efecto (espina de pescado) Gráfica mediante la cual los miembros de un equipo representan, categorizan y evalúan todos los posibles motivos de un

* *En el anteproyecto, en los pasos 10 y 11, sólo se plantean las ideas específicas para realizarlos. Hellriegel, D. y Slocon, J. (2004). Administración. Madrid: Thomson Editores.*

www.FreeLibros.me

resultado o una reacción; por lo general, se expresa como un problema para resolver. Se le conoce como diagrama de Ishikaw (Hellriegel & Slocon).

Gráficas de control Se utilizan para hacer control de calidad de procesos. Según Levin & Rubin (1996), "estas gráficas también se conocen con el nombre de diagramas de control y son de varios tipos" (p. 179)

- Diagramas $\overline{X}$ o diagramas de control para medidas de procesos.
- Diagramas *R* o diagramas de control para variabilidad de procesos.
- Diagramas ρ o diagramas de control para atributos.

Distribución de frecuencias y representaciones gráficas Según Mason y Lind (1997), "la distribución de frecuencias es el agrupamiento de datos en categorías que muestran el número de observaciones de cada categoría" (p. 24). En otras palabras, una distribución de frecuencias indica el número de veces que ocurre cada valor o dato en una tabla de resultados de un trabajo de campo.

- *Histogramas*: son medios gráficos para representación de la distribución de frecuencias.
- *Polígonos de frecuencia*: al igual que el histograma, son gráficas que permiten obtener una imagen rápida de las principales características de los datos de una distribución de frecuencias.
- *Gráficas de barras o pie (pastel)*: son formas distintas de representar los datos de una investigación.

Medidas de tendencia central

- *La media*: es la sumatoria de un conjunto de puntajes dividida por el número total de éstos.
- *La moda*: es el puntaje que ocurre con mayor frecuencia en una distribución de datos.
- *La mediana*: es el valor que divide a una distribución de frecuencias por la mitad, una vez ordenados los datos de manera ascendente o descendente.

Medidas de dispersión

- *Varianza*: es la suma de las desviaciones de la media elevadas al cuadrado, dividida entre el número de observaciones menos uno.
- *Desviaciones estándares*: es la cantidad promedio en que cada uno de los puntajes individuales varía respecto a la media del conjunto de puntajes.

Pruebas estadísticas

- *Prueba t de Student*: es un estadístico de prueba que se utiliza cuando las poblaciones son pequeñas ($n \leq 30$).
- *Prueba Z*: es una prueba de distribución normal, que tiene que ver con la probabilidad de que un puntaje dado de una medición aparezca en una distribución.

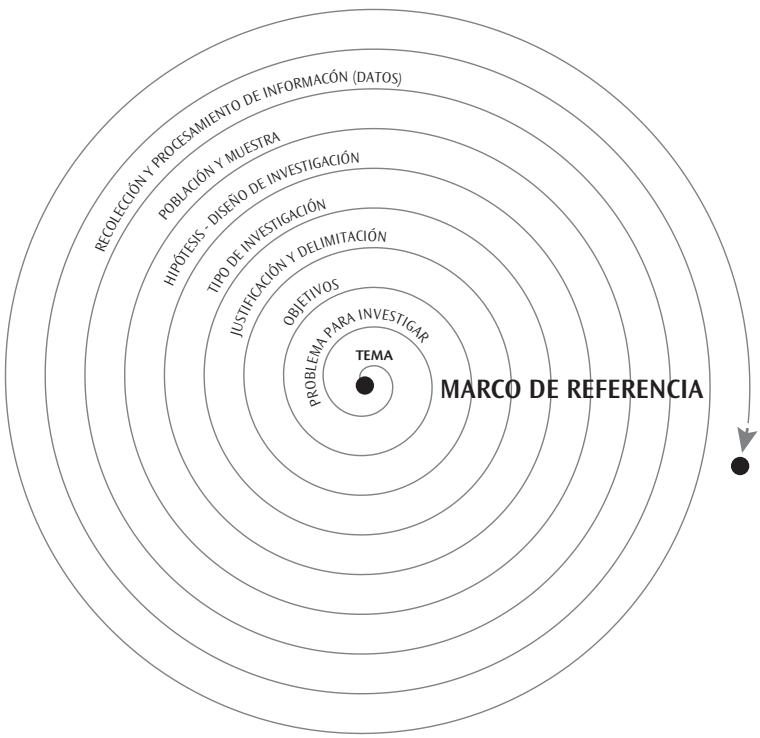
www.FreeLibros.me

- *Análisis de varianza*: es una prueba estadística para analizar si más de dos grupos difieren significativamente entre sí, en cuanto a sus medidas y varianzas.
- *Análisis de covarianza*: es una prueba que se usa para analizar la existencia o no de relación entre una variable dependiente y dos o más independientes.
- *Chi cuadrado*: es una prueba estadística que permite probar si más de dos proporciones de población pueden considerarse iguales; o, en otras palabras, permite probar si esas proporciones no presentan diferencias significativas.
- *Análisis de regresión y correlación.*
- *Análisis de regresión múltiple.*
- *Análisis de factores.*
- *Análisis multivariado de varianza (Manova).*

En la actualidad, como se ha mencionado, en un proceso de investigación científica, los análisis estadísticos se realizan mediante el uso de programas estadísticos por computador, como el Stapgraphic o el SPSS.

La Figura 7.16 ilustra la relación entre la fase de procesamiento de la información y otros aspectos de la espiral del proceso de investigación científica.

FIGURA 7.16 **El proceso de investigación. Procesamiento de la información**



www.FreeLibros.me

RESUMEN

El procesamiento de datos debe realizarse mediante el uso de herramientas estadísticas con el apoyo del computador, utilizando alguno de los programas estadísticos que hoy fácilmente se encuentran en el mercado.

Para efectuar un procesamiento de datos se deben seguir estos pasos:

- Obtención de la información de la población o muestra objeto de la investigación.
- Definición de las variables o los criterios para ordenar los datos obtenidos del trabajo de campo.
- Definición de las herramientas estadísticas y del programa de cómputo que va a utilizarse en el procesamiento de datos.
- Introducción de los datos en el computador y activar el programa para que procese la información.
- Impresión de los resultados.

El procesamiento de resultados puede efectuarse mediante:

1. Análisis de Pareto
2. Diagrama de causa/efecto
3. Gráficas de control.
4. Distribución de frecuencias y representaciones gráficas:
   - Histogramas
   - Polígonos de frecuencia
   - Gráficas de barras o *pie*
5. Medidas de tendencia central:
   - La media
   - La moda
   - La mediana
6. Medidas de dispersión:
   - Varianza
   - Desviaciones estándares
7. Pruebas estadísticas:
   - Prueba *t* de Student
   - Prueba *Z*
   - Análisis de varianza
   - Análisis de covarianza
   - Chi cuadrado
   - Análisis de regresión y correlación
   - Análisis de regresión múltiple
   - Análisis de factores
   - Análisis multivariado de varianza (Manova)
8. Otras herramientas estadísticas

www.FreeLibros.me

EJERCICIOS DE REPASO Y ANÁLISIS

1. Explique en qué consiste el procesamiento de la información en un proyecto de investigación.
2. ¿Cuáles son las principales herramientas estadísticas utilizadas en el procesamiento de los resultados?
3. ¿De qué depende el uso de una u otra herramienta estadística en el procesamiento de datos de una investigación?
4. ¿Cuáles son los pasos que deben seguirse en el procesamiento de los datos en la investigación científica?
5. Seleccione un tema de investigación de su interés del campo de su disciplina o profesión, luego describa cómo sería el procesamiento de los datos que se obtendría en el evento de realizarse la investigación del tema elegido.
6. Suponga que tiene interés en realizar una investigación tendiente a conocer las principales implicaciones del estrés causado por sobrecarga en el trabajo, en el rendimiento laboral de las personas, y ha delimitado el estudio a personal operativo del sector financiero de la ciudad QLPTY. ¿Cómo sería el procesamiento de la información en este caso?
7. Consulte informes de investigación publicados en bibliografía especializada y revise cómo se hizo el procesamiento de datos en algunos de los informes.

www.FreeLibros.me

7.12 Modelos de procesamiento de datos con el uso de herramientas estadísticas

7.12.1 Distribución de frecuencias

Una *distribución de frecuencias* es un conjunto de puntuaciones presentadas en una tabla de manera ordenada, según características definidas por el investigador.

EJEMPLO 7.26

Los datos presentados en la Tabla 7.10 son el resultado de una investigación, cuyo propósito era conocer el nivel de desempleo de una importante ciudad de su país, según el sexo y el nivel académico de las personas desempleadas.

TABLA 7.10 **E.1. Nivel de desempleo en la ciudad KIO, según sexo y nivel académico de la población**

Nivel académico del desempleado	Sexo			
	Hombres	Mujeres	Total	
			No.	%
Profesional	740	987	1 727	13,7
Técnico profesional	580	470	1 050	8,3
Bachiller	2 300	2 950	5 250	41,5
Sin bachillerato	2 190	2 420	4 610	36,5
Total	5 810	6 827	12 637	100,0
Porcentaje	46%	54%		

Interpretación: los datos de la Tabla 7.10 muestran que de un total de 12 637 personas desempleadas entrevistadas en la ciudad KIO, el mayor índice de desempleo se da entre bachilleres (41,5%), seguido de personas con nivel educativo inferior al bachillerato (36,5%) y el menor nivel de desempleo se da en el de técnicos profesionales (8,3%). En relación con el sexo, el mayor desempleo en la ciudad se presenta en las mujeres: 54% del total de los entrevistados.

7.12.2 Histogramas

Las distribuciones de frecuencia se presentan muy a menudo en forma de figuras gráficas denominadas histogramas (gráficas de barras) o en gráficas circulares.

www.FreeLibros.me

Para el caso del ejemplo de la investigación sobre el desempleo en la ciudad KIO, la gráfica de barras o histograma se presenta así:

FIGURA 7.17 **Desempleo en la ciudad de KIO, según nivel académico y sexo de las personas**

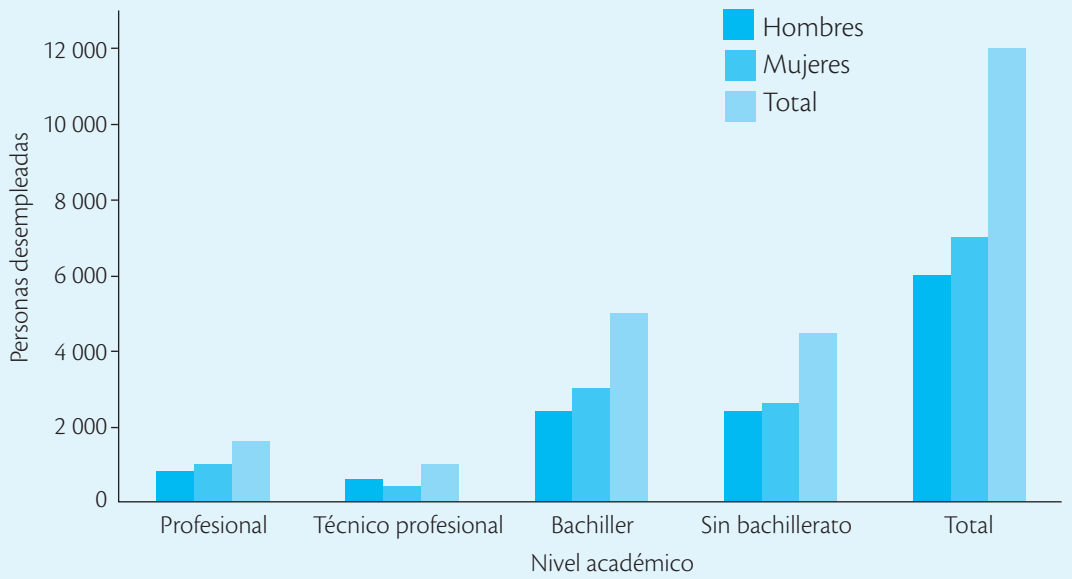
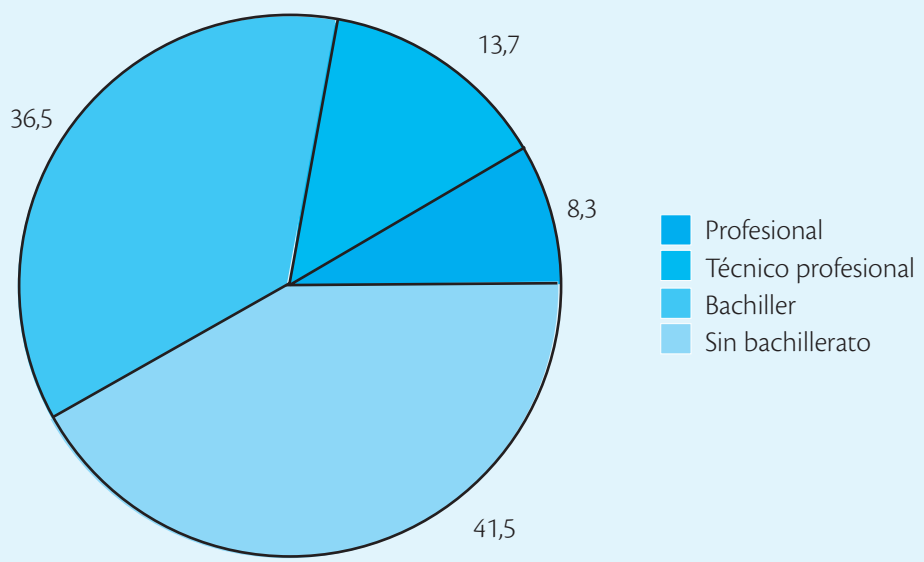


FIGURA 7.18 **Nivel académico frente al total de desempleo en la ciudad, según el nivel académico**



www.FreeLibros.me

7.12.3 Medidas de tendencia central

Las medidas de *tendencia central* son cantidades típicas o representativas de un conjunto de datos; las principales medidas son: moda, mediana, y media o promedio.

- La *moda* es la categoría o puntuación que ocurre con mayor frecuencia en un registro de datos.

EJEMPLO 7.27

El gerente de un supermercado ha realizado una investigación tendiente a medir el número de veces por mes que las mismas personas visitan su almacén y ha encontrado los siguientes resultados:

TABLA 7.11 **Visitas por mes a los almacenes TLP**

Número de personas entrevistadas	Número de visitas por mes
520	3
750	2
670	4

En este caso, la moda es 2 veces, porque es el número que más se repite, ya que 750 personas dicen visitar dos veces en el mes el supermercado TLP.

- La *mediana* es el valor que divide una distribución de frecuencias por la mitad, una vez ordenados los datos de manera ascendente o descendente.

EJEMPLO 7.28

Un investigador interesado en conocer el número promedio de consumo de litros de cerveza por persona en los habitantes de la ciudad de Building, realizó un estudio exploratorio en un expendio de cerveza, donde entrevistó a 21 personas y encontró los siguientes resultados, que aparecen en la Tabla 7.12.

TABLA 7.12 **Consumo promedio per cápita mensual de cerveza de los habitantes de Building**

Número de personas entrevistadas	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Número de litros de cerveza consumida	5	6	7	7	7	7	8	8	9	9	9	9	10	11	11	12	13	14	15	15	17

10 entrevistados — Mediana — 10 entrevistados

www.FreeLibros.me

En este caso, la mediana es el dato de la posición 11 (9 litros de cerveza) que indica que la población encuestada consume el equivalente de 9 litros de cerveza mensualmente.

- La media o el promedio es la medida de tendencia central más utilizada y se define como el promedio aritmético de una distribución. Usualmente se simboliza como $\overline{X}$, y es la suma de todos los valores de una medición dividida por el número de mediciones. La fórmula para estimar el promedio puede representarse así:

$$\overline{X} = \frac{x_1 + x_2 + x_3 + \ldots + x_n}{n} \qquad \overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Donde:

X_i : cada uno de los datos de la medición.

n : número de datos sumados..

EJEMPLO 7.29

Siguiendo el caso del consumo promedio per cápita de litros de cerveza por los habitantes de Building, el promedio es el siguiente:

$$\overline{X} = \frac{6 + 8 + 13 + 7 + 8 + 5 + 9 + 11 + 5 + 15 + 12 + 9 + 10 + 7 + 15 + 9 + 11 + 9 + 17 + 14 + 7}{21} = 9{,}95$$

Interpretación: los datos de la encuesta a 21 personas de Building indican que el promedio per cápita de consumo mensual de cerveza en esta ciudad es de 10 litros de cerveza, aproximadamente.

7.12.4 Medidas de dispersión

Las medidas de *dispersión* que se emplean con mayor frecuencia incluyen la desviación estándar, la varianza y el rango. Las medidas de dispersión indican cuán "dispersos" o separados se encuentran los datos, respecto a un valor central.

- **Rango o intervalo**

El *rango* es la más elemental de las medidas de tendencia central; consiste simplemente en la distancia entre los dos valores más extremos de una medición.

EJEMPLO 7.30

Siguiendo el caso del consumo promedio per cápita mensual de litros de cerveza por los habitantes de Building, el rango se calcula así:

www.FreeLibros.me

Se toma el valor mayor de litros de cerveza consumido y se le resta el valor menor de todos los entrevistados; en este caso, los valores son:

17 litros de cerveza - 6 litros de cerveza = 11 litros de cerveza

Esto significa que existe una diferencia de 11 litros de cerveza, entre la persona que más la consume al mes y la que menos consume en esa ciudad.

- **Desviación estándar**

La *desviación estándar* es el promedio de desviaciones o dispersiones de las puntuaciones respecto a la media o al promedio. Es decir, la desviación estándar permite medir el grado de homogeneidad o heterogeneidad de los datos de la población objeto de medición. Cuanto mayor sea la dispersión de los datos respecto a la media, mayor será la desviación estándar, lo cual significa mayor heterogeneidad en las mediciones.

La fórmula para calcular la desviación estándar de una muestra de observaciones es:

$$S = \sqrt{\frac{\sum_{i=1}^{n}\left(X_1 - \overline{X}\right)^2}{n-1}}$$

Donde:

S : desviación estándar.

X_i : valor de cada una de las mediciones.

$\overline{X}$: media de la muestra de mediciones.

n : tamaño de la muestra.

EJEMPLO 7.31

Continuando con el estudio sobre consumo promedio per cápita mensual de litros de cerveza por los habitantes de Building, la desviación estándar se calcula así:

TABLA 7.13 **Consumo promedio mensual de cerveza en la ciudad de Building**

No. de encuestados	Litros de cerveza X_i	$(X_i - X)^2$
1	6	15,21
2	8	3,61
3	13	9,61
4	7	8,41
5	8	3,61

Continúa Tabla 7.13

www.FreeLibros.me

No. de encuestados	Litros de cerveza X_i	$(X_i - X)^2$
6	5	24,01
7	9	0,81
8	11	1,21
9	7	8,41
10	15	26,01
11	12	4,41
12	9	0,81
13	10	0,01
14	7	8,41
15	15	26,01
16	9	0,81
17	11	1,21
18	9	0,81
19	17	50,41
20	14	16,81
21	7	8,41
Suma	$\overline{X} = \sum x_i/n$	219,01

$$S = \sqrt{\frac{219{,}01}{20}} = 3{,}30 \text{ litros de cerveza.}$$

Interpretación: el consumo promedio per cápita mensual de litros de cerveza de la muestra de la población de Building se dispersa en promedio, respecto a la media, en 3,30 litros de cerveza.

EJEMPLO 7.32

El siguiente es otro ejemplo de interpretación del valor de la desviación estándar.

Suponga que usted realizó una investigación cuyo propósito fue conocer el número de quejas promedio mensuales, efectuadas por los clientes sobre la calidad del servicio prestado en un hotel localizado en la ciudad de Ameranta. Para la muestra de los últimos seis meses, usted obtuvo una media de $\overline{X} = 79$ quejas por mes, y una desviación estándar de $S = 16$ quejas por mes.

La interpretación es que el número mensual de quejas de la muestra se dispersan en promedio, respecto del promedio 16 quejas.

www.FreeLibros.me

- **Varianza**

Se define la *varianza* como la desviación estándar elevada al cuadrado.

7.12.5 Prueba *Z*

La prueba *Z* es una prueba de distribución normal que tiene como finalidad comparar los puntajes de distribuciones que son diferentes entre sí.

Para Salkind (1998), "los valores de los *puntajes Z* tienen que ver con la probabilidad de que un puntaje dado aparezca en una distribución" (p. 177).

El valor de *Z* es el resultado de dividir la desviación de un puntaje individual respecto a la media, entre la desviación estándar, utilizando la fórmula siguiente para calcular la prueba *Z*.

$$Z = \frac{x_1 - \overline{X}}{S}$$

Donde:

X_i : puntuación o valor de medición para transformar en unidades de desviación estándar.

$\overline{X}$: media de la distribución o el promedio de las mediciones.

S : desviación estándar de las mediciones o los datos obtenidos.

Z : puntuación de la medición transformada a unidades de desviación estándar.

La estandarización de valores permite comparar puntuaciones de dos distribuciones diferentes.

A continuación, un ejemplo de aplicación de las puntuaciones *Z* o estandarización de mediciones.

EJEMPLO 7.33

Suponga que va a compararse una distribución, obtenida en la medición del valor promedio mensual de gastos per cápita en alimentación por los habitantes de una ciudad *A*, con una distribución de medición de la misma variable en la ciudad *B*.

Suponga también que un habitante de la ciudad *A* gasta en promedio al mes, para el rubro alimentación, $270 000 (la media de gasto en alimentación para esta ciudad es de $305 000 y la desviación estándar de $39 000); a la ciudad *B*, otro habitante destina en alimentación $295 000 por mes (el promedio de gasto mensual per cápita de los habitantes de la ciudad *B* es $285 000 con una desviación estándar de $30 000).

¿Está en mejores condiciones de gasto en alimentación el habitante de la ciudad *B* respecto al de la ciudad *A*?

www.FreeLibros.me

Entonces:

$$Z_A = \frac{270\,000 - 305\,000}{39\,000} = -0{,}089$$

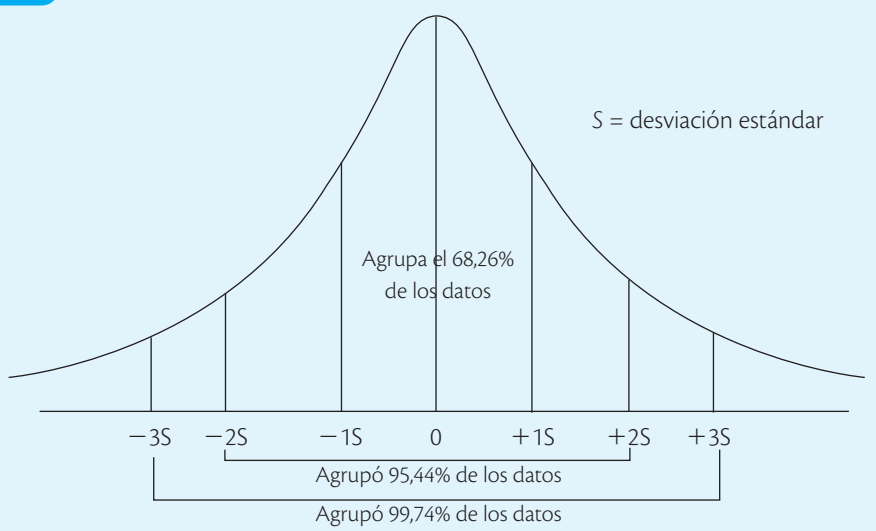
$$Z_B = \frac{295\,000 - 285\,000}{30\,000} = -0{,}33$$

Como los valores de Z tanto para la ciudad A como para la ciudad B son inferiores a 1,96, entonces no existen diferencias significativas entre el gasto del habitante de una ciudad con respecto a la otra, porque ambos están alrededor del valor $Z = 1{,}96$.

Aparecen diferencias cuando uno de los valores Z es inferior a 1,96 y el otro superior.

En la distribución de puntuaciones Z, estos valores se ubican así:

FIGURA 7.19 **Distribución de puntuaciones Z**



Prueba de hipótesis

Recuérdese que una hipótesis es una suposición respecto al problema de investigación, y lo que se hace en la prueba de hipótesis es determinar si la proposición es consistente con los datos obtenidos una vez realizada la investigación. Si la hipótesis o proposición no es consistente con los datos obtenidos, se rechaza la hipótesis.

Pasos para probar las hipótesis

1. **Formular la hipótesis.** Las hipótesis se plantean de dos manera fundamentales: la hipótesis nula H_o y la hipótesis alterna H_A. Por ejemplo, el gerente del Banco *XPQ* considera que sus procedimientos operativos garantizan que el cliente promedio sólo tiene que esperar siete minutos en la fila de los cajeros antes de ser atendido.

www.FreeLibros.me

La hipótesis nula y la hipótesis alterna se formulan de la siguiente manera:

Hipótesis nula H_o: tiempo promedio de espera en la fila para ser atendido = 7 minutos.

Hipótesis alterna H_A: tiempo medio de espera en la fila para ser atendido ≠ 7 minutos.

2. **Elegir la prueba estadística adecuada.** Es muy importante saber que para la prueba de hipótesis existen varias pruebas estadísticas, por lo que el investigador debe elegir la apropiada teniendo en cuenta las características del caso que va a investigar.
3. **Definir el nivel de significancia.** Por ejemplo: $\alpha = 0{,}05$.
4. **Recolectar los datos con una muestra representativa.** En el caso del estudio del tiempo de espera por los clientes del Banco *XPQ*, se lleva a cabo una investigación con base en la observación en momentos aleatorios de 850 clientes en oficinas del banco elegidas aleatoriamente. El tiempo promedio de espera de los clientes fue de 8,2 minutos y la desviación estándar de 2,6 minutos.
5. **Estimar la desviación estándar de la distribución muestral de la media.** Se utiliza la siguiente fórmula:

$$S_x = \frac{S}{\sqrt{n}}$$

Donde:

S_x : desviación estándar de la distribución muestral de la media.

S : desviación estándar de la muestra.

n : tamaño de la muestra.

Entonces:

$$S_x = \frac{2,6}{\sqrt{850}} = 0{,}089$$

Para transformar la media de la muestra en una puntuación *Z*, en el contexto de la distribución muestral:

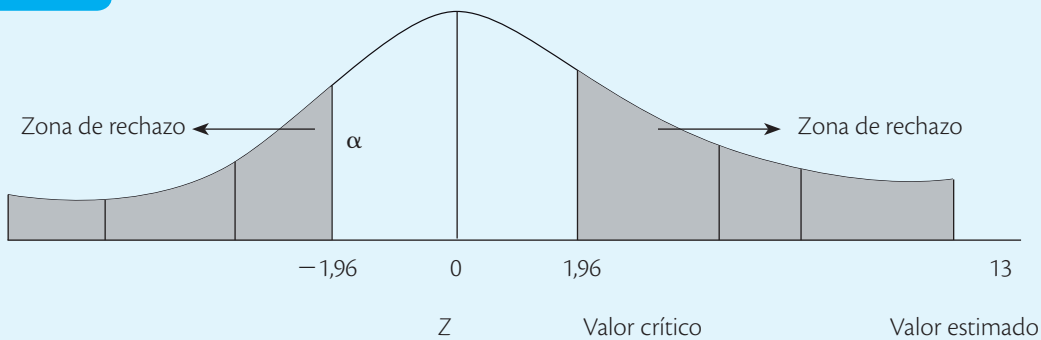
$$Z = \frac{x - \overline{X}}{S_X} = \frac{8{,}2 - 7{,}0}{0{,}089} = 13$$

Entonces, se compara el valor calculado (en este caso $Z = 13$) con el valor crítico de la prueba estadística (tomada de la tabla para valores *Z*), según el nivel de significación elegido.

Si el valor calculado ($Z = 13$) es menor que el valor crítico ($Z = 1{,}96$ con nivel de significancia $\alpha = 0{,}05$), no se rechaza la hipótesis nula. Si el valor calculado es mayor que el valor crítico, se rechaza la hipótesis nula.

www.FreeLibros.me

FIGURA 7.20 **Valor crítico y valor estimado**



6. **Decisión estadística.** Para el caso referido, como el valor calculado $Z = 13$ es mayor que el valor crítico (*véase* tabla Z) 1,96, entonces se rechaza la hipótesis nula.
7. **Conclusión.** Con un nivel de significancia del 0,05 ó 95% de confianza, no es adecuado afirmar que el tiempo promedio de espera en la fila por un cliente del Banco *XPQ* es de 7 minutos.

7.12.6 Prueba *t*

La prueba *t* de *Student* es una prueba estadística para evaluar hipótesis en torno a una media, cuando los tamaños de la muestra n son menores de 30 mediciones ($n < 30$), y se quiere saber si existe diferencia significativa entre la media de la muestra y la media poblacional.

El valor *t* se obtiene mediante la siguiente fórmula:

$$t = \frac{(\textit{media de la muestra}) - (\textit{media poblacional según hipótesis nula})}{\textit{error estándar estimado de la media}} = \frac{\overline{X} - \mu}{S_x}$$

EJEMPLO 7.34

Prueba de hipótesis con la prueba *t*: suponga que el gerente comercial de Editorial *CABW* estima que es necesario vender 750 libros de un determinado título por semestre, en cada una de las nueve ciudades del país, para que la venta de los libros sea rentable para la editorial.

El gerente quiere probar la hipótesis de que las ventas por ciudad son superiores a los 750 libros por semestre.

1. Para ello se formulan las siguientes hipótesis:

H_o: las ventas promedio de libros por semestre, por ciudad es menor o igual que 750 libros.

H_A: las ventas promedio de libros por semestre por ciudad es mayor que 750 libros.

www.FreeLibros.me

Planteadas las hipótesis, el siguiente paso es establecer el nivel de significancia (valor de α) permitido. Para $\alpha = 0{,}05$, el valor tabulado de t (valor crítico) con 8 grados de libertad ($n - 1$, donde n es el número de ciudades), $t = 1{,}86$ (*véase* tabla t para 8 grados de libertad).

2. Para obtener la información se efectuó el estimativo en cada una de las nueve ciudades y se encontró que las ventas promedio por semestre fueron de 809 libros con una desviación estándar de 118 libros.
3. Con los datos anteriores se calcula el error estándar de la media S_x, con la siguiente fórmula:

$$S_x = \frac{S}{\sqrt{n}} = \frac{118}{\sqrt{9}} = 39{,}3$$

4. Luego se calcula el valor estadístico de la prueba t:

$$t = \frac{(\textit{media de la muestra}) - (\textit{media poblacional según hipótesis nula})}{\textit{error estándar estimado de la media}} = \frac{\overline{X} - \mu}{S_x}$$

$$t = \frac{809 - 750}{39{,}3} = 1{,}50$$

Para saber si se rechaza o no la hipótesis nula, es necesario comparar el valor t estimado ($t = 1{,}50$) con el valor t crítico o valor en tabla ($t = 1{,}86$).

5. **Decisión**. Como el valor t estimado ($t = 1{,}50$) es menor que el valor t crítico ($t = 1{,}86$), entonces no se rechaza la hipótesis nula.
6. **Conclusión**. Aunque las ventas medias por semestre ($X = 809$ libros) son superiores a 750 libros, la diferencia basada en la muestra de nueve ciudades no es estadísticamente significativa. Según estos datos, la decisión de vender el volumen estimado de libros no es efectiva, lo cual indica que la venta de libros puede ser inferior a 750 libros semestrales por ciudad.

7.12.7 Hipótesis sobre proporciones

El gerente de uno de los hoteles importantes de la ciudad *CTMC* llevó a cabo una encuesta entre 450 turistas que visitaban la ciudad, y encontró que 73% de los encuestados gastaban valores superiores a US$3 500 al año en sus visitas a la ciudad, y que estarían decididos a utilizar los servicios del hotel. La compañía está interesada en desarrollar un paquete especial de servicios para este grupo de turistas si los datos son reales.

El hotel desarrollaría su plan si el porcentaje verdadero de turistas con gastos superiores a los US$3 500 es mayor que 60%.

1. Se plantea la hipótesis de trabajo:
   Hipótesis nula $H_o = P \leq 60\%$ o $H_o \leq 0{,}60$.
   Hipótesis alterna $H_A = P > 60\%$ o $H_A > 0{,}60$.
   Donde P = proporción de turistas con gastos superiores a US$3 500 al año en sus visitas a la ciudad *CTMC*.

www.FreeLibros.me

2. Se establece el nivel de significancia (valor α) permitido. Para $\alpha = 0{,}05$, el valor tabulado de Z (valor crítico) = 1,64 (consultar tabla para valores Z, una sola cola).
3. Se calcula el error estándar estimado empleando el valor P especificado en la hipótesis nula.

$$S_x = \sqrt{\frac{P(1-P)}{n}} = \sqrt{\frac{(0{,}60)(0{,}40)}{450}} = 0{,}02309$$

4. Se calcula la prueba estadística como sigue:

$$Z = \frac{\hat{P} - P}{S_p} = \frac{0{,}73 - 0{,}60}{0{,}02309} = 5{,}6301$$

5. **Decisión**. Como el valor calculado para Z = 5,63 es mayor que el valor crítico de Z (1,64), se rechaza la hipótesis nula.
6. **Conclusión**. Por los datos obtenidos en el estudio, se concluye que con un nivel de confianza de 95% (1 - $\alpha = 0{,}95$), más del 60% de los turistas de la ciudad *CTCM* realizan gastos superiores a US$3 500 al año en sus viajes a esta ciudad. Por tanto, la gerencia puede introducir el paquete de servicios en su hotel y dirigirlos a este grupo de turistas.

7.12.8 Prueba chi cuadrado para prueba de hipótesis

La prueba de chi cuadrado (X^2) permite determinar si el patrón de frecuencia observado corresponde o se ajusta al patrón esperado; también sirve para evaluar hipótesis acerca de la relación entre dos variables categóricas.

EJEMPLO 7.35 **Prueba de chi cuadrado (X^2) para una muestra única**

Suponga que el director del servicio de salud necesita probar la eficacia de tres campañas de vacunación a la población infantil en su ciudad, en los últimos tres años. El director desea medir el efecto de cada campaña en los padres de familia que lleven a sus niños a los centros de salud que prestan el servicio de vacunación durante el tiempo de la respectiva campaña.

TABLA 7.14 **Atención de niños en vacunación en los últimos tres años en la ciudad**

Año de campaña	Número de niños atendidos en mes de campaña	Mes de campaña
2010	11 700	Julio
2011	12 300	Mayo
2012	10 900	Septiembre
Total	34 950	

www.FreeLibros.me

El director necesita saber si el número de niños atendidos durante el período en que estuvo vigente cada campaña difiere en forma significativa. Este problema se resuelve aplicando la prueba de chi cuadrado (X^2) de la siguiente forma:

1. Se plantean la hipótesis nula y la hipótesis alterna:

   Hipótesis nula H_o = no existen diferencias significativas en el número de infantes atendidos en las últimas tres campañas de vacunación infantil.

   Hipótesis alterna H_A = existen diferencias significativas en la cantidad de niños atendidos en las campañas de vacunación en los últimos tres años.

2. Se determina el número de niños atendidos en cada año de campaña, si la hipótesis nula fuera correcta (E_i). Por tanto, se espera que haya un número igual de infantes atendidos en el curso de cada campaña. El número esperado se calcula sumando el total de infantes atendidos en las tres campañas y dividiendo este valor sobre el número de campañas (3).

$$E_i = \frac{T_A}{n} = \frac{\textit{Total infantes atendidos}}{\textit{Número de campañas}} = \frac{34\,950}{3} = 11\,650 \text{ niños(as)}$$

3. Se calcula el valor de X^2 mediante la ecuación

$$X_2 = \sum \frac{(O - E_i)^2}{E_i}$$

Donde:

O_i : número observado en la iésima categoría o campaña.

E_i : número esperado en la iésima categoría.

K_K : número de categorías, en este caso 3 (campañas).

$$X^2 = \frac{(11\,700 - 11\,650)^2 + (12\,350 - 11\,650)^2 + (10\,900 - 11\,650)^2}{11\,650} = 90{,}55$$

4. Se escoge el nivel de significancia permitido α. Para el nivel de significancia $\alpha = 0{,}05$, el valor tabulado de X^2 con dos grados de libertad (k - 1) es 5,99 (*véase* tabla X^2).

5. **Decisión**. Se compara el valor calculado X^2 con el valor crítico (valor en tabla) de X^2. Como el valor calculado para X^2 (90,55) es mayor que el valor tabulado (5,99), entonces se rechaza la hipótesis nula, que afirma que no existe diferencia significativa en el número de niños atendidos por campaña de vacunación.

6. **Conclusión**. Se concluye que con 95% de confianza en los datos obtenidos en las campañas de vacunación infantil, en los últimos tres años en la ciudad objeto del estudio, existen diferencias significativas en el número de niños atendidos entre cada una de las distintas campañas de vacunación.

www.FreeLibros.me

Análisis de regresión y correlación

El análisis de regresión y correlación es un método estadístico utilizado para calcular la relación entre dos o más variables y su grado de relación.

1. Análisis de regresión lineal

Para ejemplificar este método de análisis de datos, se plantea el ejemplo 7.36.

EJEMPLO 7.36

El gerente de un prestigioso restaurante de la ciudad APQ realizó una investigación en su negocio con el propósito de analizar la relación entre el número de clientes atendidos y el número de quejas recibidas en cada uno de los últimos seis meses del año en curso.

Los datos de clientes y de quejas de los seis meses objeto del estudio se muestran en la Tabla 7.15.

TABLA 7.15 **Número de clientes atendidos y quejas recibidas por el hotel**

Mes de actividad	Número de clientes atendidos por mes (miles)	Número de quejas recibidas por mes
1	6,6	75
2	5,3	69
3	4,9	71
4	7,4	86
5	7,1	78
6	5,6	73

1. El primer paso en el análisis de regresión es recurrir al método de mínimos cuadrados, que es una técnica matemática que permite construir con los datos X y Y la línea que representa mejor la relación entre las dos variables.

   La ecuación general del método de los mínimos cuadrados que se emplea en el análisis de regresión es:

   $$Y = a + bx$$

   Donde:

   Y : variable dependiente (en este caso, número de quejas).
   a : intersección estimada de la línea de regresión con el eje Y.
   b : pendiente estimada de la línea de regresión: coeficiente de regresión.
   x : variable independiente, número de clientes que visitan el restaurante.

www.FreeLibros.me

2. El siguiente paso en el análisis de regresión es calcular los valores apropiados de a y b para sustituirlos en la ecuación de mínimos cuadrados; entonces:

$$a = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{(n)(\sum x^2) - (\sum x)^2}$$

$$b = \frac{(n)(\sum xy) - (\sum x)\sum (y)}{(n)\sum (x^2) - (\sum x)^2}$$

Donde:

n : número de datos.
$\sum X$: suma de los valores X.
$\sum Y$: suma de los valores Y.
$\sum X^2$: suma de los valores de X^2.
$\sum XY$: suma de los productos de X por Y.
a : intersección de la línea de regresión con el eje Y.
b : pendiente estimada de la línea de regresión: coeficiente de regresión.

Cálculo de la ecuación

TABLA 7.16 **Cálculo de regresión lineal para el número de clientes atendidos y quejas recibidas para el hotel**

Mes de actividad	X	Y	X^2	XY	Y^2
1	6,6	75	43,56	495,0	5 625
2	5,3	69	28,09	365,7	4 761
3	4,9	71	24,01	347,9	5 041
4	7,4	86	54,76	636,4	7 396
5	7,1	78	50,41	553,8	6 084
6	5,6	73	31,36	408,8	5 329
$n = 6$	36,9	452	232,19	2 807,6	34 236

3. Se remplazan los datos en la ecuación de mínimos cuadrados y se obtiene la ecuación de regresión lineal:

$$Y = a + bx$$

Donde:

$$a = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{(n)(\sum x^2) - (\sum x)^2} = \frac{(452)(232{,}19) - (36{,}9)(2.807{,}6)}{(6)\,232{,}19 - (36{,}9)^2} = 42{,}8$$

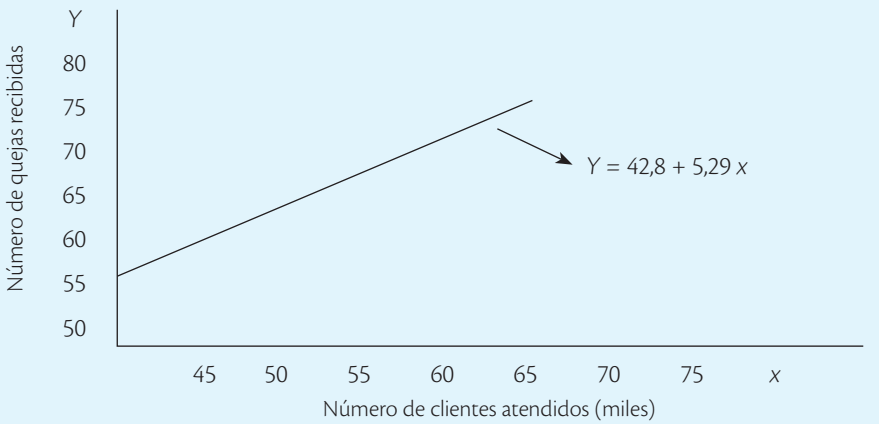
$$b = \frac{(n)\sum (xy) - (\sum x)\sum (y)}{(n)\sum (x^2) - (\sum x)^2} = \frac{(6)\,2.807{,}6 - (36{,}9)(452)}{(6)\,232{,}19 - (36{,}9)^2} = 5{,}29$$

$$Y = 42{,}8 + 5{,}29\,x$$

www.FreeLibros.me

La línea que representa mejor la relación entre las dos variables para esta ecuación es la que muestra la Figura 7.21.

FIGURA 7.21 **Relación entre clientes atendidos y quejas recibidas por mes**



Los datos de la figura indican que existe una relación directa (positiva) entre el número de clientes atendidos y el número de quejas recibidas, es decir, a medida que aumenta el número de clientes atendidos por el restaurante, aumenta el número de quejas recibidas por el servicio prestado a sus clientes.

2. Coeficiente de correlación

El *análisis de correlación* sirve para medir la fuerza o el grado de correlación entre las variables objeto de estudio en el análisis de regresión.

La ecuación para medir el *coeficiente de correlación* es la siguiente:

$$r = \frac{(n)(\sum xy) - (\sum x)\sum (y)}{\sqrt{[(n)\sum x^2 - (\sum x)^2]\ [(n)\sum y^2 - (\sum y)^2]}}$$

Donde:

r : coeficiente de correlación.

El coeficiente de correlación es un número que, en determinado conjunto de datos, se encontrará entre -1 y +1 y que indica:

a. Dirección de la correlación. Si es positiva, significa que X y Y aumentarán y disminuirán simultáneamente; por ejemplo, si X se incrementa, Y tenderá entonces a aumentar también. Si r es negativa, las variables tenderán a moverse en direcciones opuestas; por ejemplo, si X se incrementa, Y tenderá a decrecer, o viceversa.

www.FreeLibros.me

b. Fuerza de la relación. Cuanto mayor sea el valor absoluto de r, más estrecha será la relación de las dos variables y mejor ajustará los datos el diagrama de dispersión de la ecuación de mínimos cuadrados. En los extremos (esto es, si *r* es +1 ó -1), la ecuación contendrá todos los puntos de datos. Esto es, cuanto más se aproxime *r a* +1, mayor será la relación directa entre las variables; y cuanto más se aproxime *r a* -1, más inverso será el grado de relación entre las variables.

Cuando $r = 0$, no hay relación lineal entre las variables.

Se calcula entonces el valor de *r* para el caso de la relación entre número de clientes atendidos y número de quejas recibidas por el restaurante objeto del estudio:

$$r = \frac{(6)\,2\,807{,}6 - (36{,}9)\,(452)}{\sqrt{[(6)\,(232{,}19 - (36{,}9)^2]\,[(6)\,(34{,}236) - (452)^2]}} = \frac{166{,}8}{187{,}24} = 0{,}89$$

Interpretación: como el valor de *r* calculado (0,89) es positivo, entonces la relación entre el número de clientes atendidos y el número de quejas recibidas es directa, es decir, cuando aumenta el número de clientes atendidos por el restaurante aumenta el número de quejas recibidas.

También, como el valor de *r* (0,89) se acerca al valor +1, significa que hay una relación muy estrecha entre el número de clientes y el número de quejas en el restaurante.

3. Coeficiente de determinación

El *coeficiente de determinación* es el cuadrado del coeficiente de correlación y asume un significado especial porque su valor representa la proporción de la variación de *Y* que se explica por la variable independiente *X* mediante la ecuación de regresión.

Siguiendo el caso del restaurante:

$$r^2 = r \times r = 0{,}89 \times 0{,}89 = 0{,}7921$$

Donde:

r : coeficiente de correlación.

r^2 : coeficiente de determinación.

Interpretación del valor 0,7921

El valor 0,7921 significa que 79,21% de la varianza en las quejas recibidas por el restaurante se explican por los cambios en el número de clientes atendidos mensualmente por el restaurante. La otra parte de la varianza en las quejas (21%) se debe a otros factores diferentes del cambio en el número de clientes atendidos.

Otra fórmula para calcular r^2 es:

$$r^2 = \frac{\textit{Varianza en Y explicada por X}}{\textit{Variación total en Y}}$$

www.FreeLibros.me

7.13 Análisis de resultados. Discusión

Una vez procesados los datos por medios estadísticos, se obtienen unos resultados que deben ser analizados e interpretados o discutidos.

El análisis de resultados consiste en interpretar los hallazgos relacionados con el problema de investigación, los objetivos propuestos, la hipótesis y/o preguntas formuladas, y las teorías o presupuestos planteados en el marco teórico, con la finalidad de evaluar si confirman las teorías o no, y se generan debates con la teoría ya existente.

En este análisis deben mostrarse las implicaciones de la investigación realizada para futuras teorías e investigaciones.

En términos generales, en el análisis también debe indicarse si el estudio respondió o no a las hipótesis o preguntas planteadas para desarrollar los objetivos del estudio. El hecho de no encontrar respaldo a la hipótesis o preguntas de investigación no debe ser motivo para considerar que el estudio fracasó; éste podría ser un excelente pretexto para iniciar un nuevo estudio que permita corroborar o contrastar los resultados encontrados.

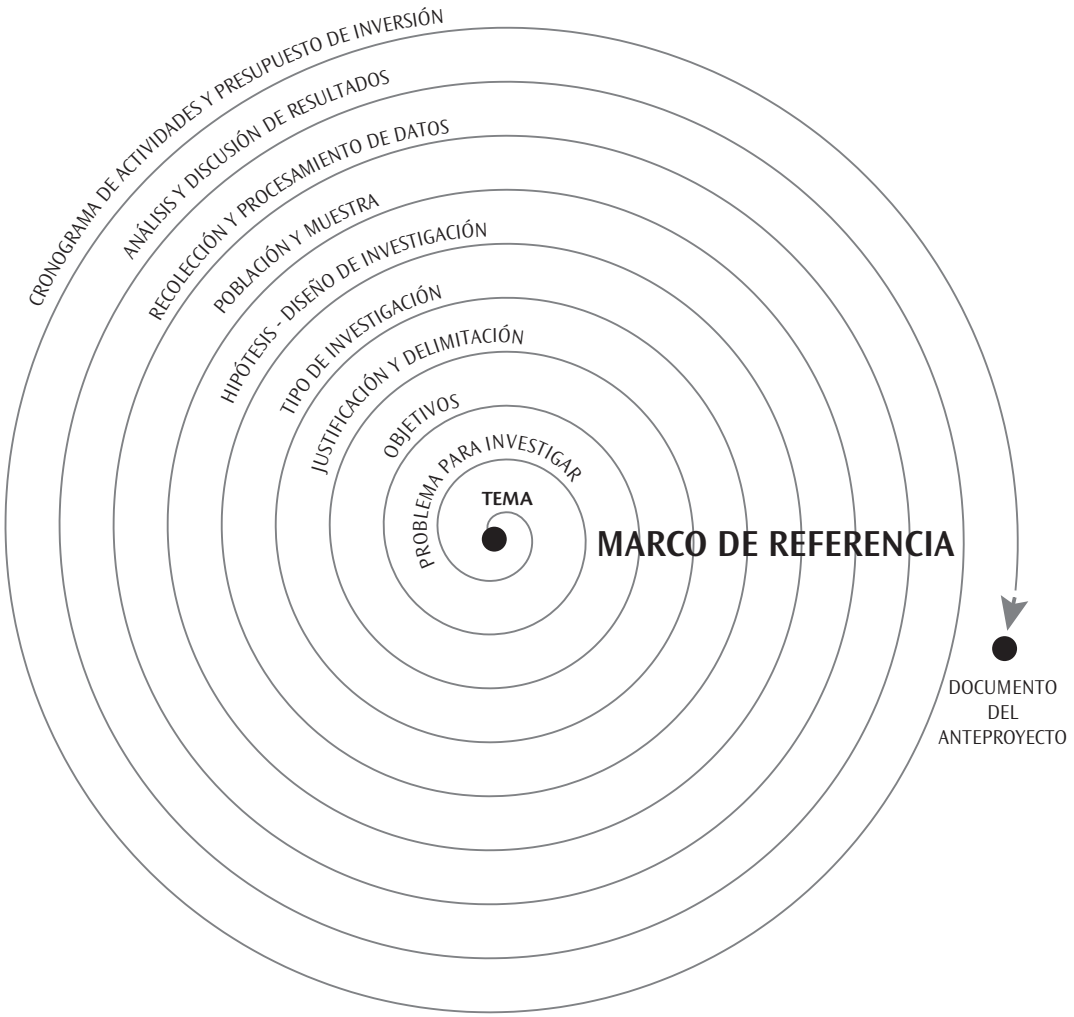
El análisis y la discusión de los resultados es el aspecto más importante que se va a tener en cuenta en toda investigación; sobre él deben hacer énfasis los jurados evaluadores del informe final presentado por los investigadores.

Nota: en el documento del *anteproyecto* se debe mencionar cómo se harán el análisis y la discusión de los resultados, en tanto que en el documento del *informe final* hay que mostrar el análisis que se hizo de los resultados de la investigación y la discusión a que dieron lugar los resultados, al contrastarlos o relacionarlos con la información presentada en el marco teórico que fundamenta la investigación.

La figura 7. 22 ilustra la relación entre la forma como se realizará el análisis de resultados y los demás componentes de la espiral del proceso de investigación científica.

www.FreeLibros.me

FIGURA 7.22 **El proceso de investigación. Análisis y discusión de los resultados**



www.FreeLibros.me

RESUMEN

El análisis y la discusión de resultados obtenidos del procesamiento de la información consiste en interpretar los hallazgos relacionados con el problema de investigación, los objetivos propuestos, la hipótesis y/o preguntas formuladas, y las teorías o los presupuestos planteados en el marco teórico, con la finalidad de evaluar si confirman las teorías o no, y se generan debates con la teoría ya existente. En este análisis y discusión deben mostrarse las implicaciones de la investigación realizada para futuras teorías e investigaciones.

EJERCICIOS DE REPASO Y ANÁLISIS

1. ¿En qué consisten el análisis y la discusión de los resultados en un proyecto de investigación?
2. ¿Cuál es la diferencia entre descripción, análisis y discusión de los resultados de una investigación?
3. Explique la relación entre el problema de investigación, los objetivos, el marco teórico y el análisis y la discusión de los resultados.
4. Busque en libros o revistas dos artículos de investigación y evalúe la relación entre los objetivos propuestos en el estudio y el análisis y la discusión de resultados.
5. Consulte revistas o material bibliográfico especializado en publicación de informes de investigación en el campo de su disciplina, lo mismo que artículos de informes de investigación, especialmente en lo relacionado con el análisis y la discusión de resultados que se hayan hecho en los informes revisados. Comparta la información obtenida con sus compañeros de clases.
6. Suponga que está interesado en un tema de investigación sobre el cual se han publicado resultados recientes. Si quiere replicar dicho estudio, ¿en qué consistiría la discusión de resultados para el caso?

www.FreeLibros.me

7.14. Cronograma de actividades, presupuesto y bibliografía

Toda investigación en su fase de anteproyecto requiere la elaboración de un cronograma y de un presupuesto de inversión para administrar la dedicación de tiempo y los recursos financieros requeridos para el desarrollo de las diferentes actividades previstas para la investigación que se va a realizar. En este sentido, al cronograma de actividades y al presupuesto de inversión se les denomina *aspectos administrativos de la investigación*.

7.14.1 Cronograma de actividades

En un documento de anteproyecto de investigación, el cronograma consiste en describir y programar las actividades requeridas para desarrollar la investigación que se va a realizar según la secuencia y su respectiva demanda de tiempo, por tal razón, para elaborar el cronograma de actividades, lo primero que se requiere es tener claridad sobre cuáles son las actividades que se deben realizar para desarrollar la investigación prevista según los objetivos de la misma (y la hipótesis que se debe probar cuando sea necesario).

También, el cronograma debe elaborarse de acuerdo con la disponibilidad del tiempo y experiencia en investigación de las personas que conformarán el equipo responsable de realizar la investigación, los recursos financieros y el tiempo total en el que se espera deben concluir la misma. Tanto el proyecto en su conjunto como cada actividad deben tener una precisa estimación de su duración para su desarrollo, previendo cierto margen de tolerancia por circunstancias de contingencias o imprevistos que puedan presentarse durante el tiempo previsto para el desarrollo de la respectiva investigación.

La representación del cronograma suele hacerse mediante diagramas que permitan visualizar de forma esquemática cada una de las actividades que se deben realizar para desarrollar la investigación, las secuencias de las mismas y su duración. Los diagramas más utilizados para la representación de los cronogramas de actividades en la investigación son las gráficas de Gantt, que suelen usarse para cronogramas de proyectos sencillos, mientras que para proyectos de mayor complejidad suelen utilizarse diagramas de redes como el PERT, el CPM y diagramas de programación asistidos por computador.

Dado que la mayoría de los proyectos de trabajo de grado en los programas de pregrado (y que son el principal objetivo de este libro) son considerados de poca complejidad en cuanto a su programación de actividades, entonces, a continuación sólo se hace referencia a la gráfica de Gantt como diagrama para representar el cronograma de actividades de un proyecto de investigación que se va a realizar. De forma sencilla, la gráfica de Gantt consta de dos columnas, en la primera, denominada actividades, muestra las diferentes actividades

www.FreeLibros.me

que se deben realizar en el desarrollo de la investigación, enlistadas de forma secuencial; en la segunda columna, denominada duración (semanas, meses y años), y mediante barras horizontales, se ilustra la duración de cada una de las actividades enlistadas en la primera columna.

La Tabla 7.17 ilustra un ejemplo general de un cronograma de actividades para un proyecto de investigación con utilización de la gráfica de Gantt.

TABLA 7.17 Cronograma de actividades

Actividades	Duración (meses, inicio 25 de agosto de 2010)										
	1	2	3	4	5	6	7	8	9	10	11
Ajuste a la propuesta según conceptos de revisores.											
Presentación del proyecto a la población objeto del estudio y diseño del cuestionario de encuesta.											
Revisión y ajuste de instrumento de encuesta.											
Desarrollo del marco teórico.											
Recolección de la información (trabajo de campo).											
Procesamiento de datos.											
Análisis de resultados.											
Informe final.											

La Tabla 7.17 ilustra las actividades previstas (primera columna) para el desarrollo general de un trabajo de grado en el pregrado, y la duración del proyecto en su conjunto (11 meses), así como la duración de cada actividad (segunda columna). Al respecto, es importante tener en cuenta que cada trabajo de grado o proyecto de investigación ha de considerar sus particularidades y las actividades serán las propias así como la duración del mismo según los criterios definidos en párrafos anteriores de esta misma sección.

Nota: en el CD adjunto a este libro, usted puede encontrar otros ejemplos de cronogramas de actividades correspondientes a los ejemplos de anteproyectos de investigación.

7.14.2 Presupuesto de inversión

Es indudable que toda investigación, por sencilla que sea, requiere el compromiso o la disponibilidad de recursos financieros para su desarrollo, por esta razón es importante elaborar el presupuesto requerido para llevar a cabo la investigación.

www.FreeLibros.me

En general, el presupuesto debe mostrar el costo total de realizar la investigación, pero también el costo detallado por los honorarios del equipo de personas responsable de la investigación, de los materiales y equipos para el desarrollo de las actividades que se van a ejecutar, del material bibliográfico y demás costos en que se incurre para el desarrollo de la investigación.

Igualmente, en el presupuesto se debe detallar la fuentes (es decir, de dónde se obtendrá el dinero) para financiar el desarrollo del respectivo proyecto. Las fuentes de los recursos financieros en el caso de los trabajos de grado suelen ser los aportes de los padres de familia o del propio estudiante; sin embargo, dependiendo del tema de investigación se pueden obtener recursos financieros de entidades interesadas en el tema, razón por la cual es importante presentar el proyecto de investigación a aquellas entidades, organizaciones o empresas que pueden llegar a tener interés en el tema por su misión institucional.

Los siguientes son algunos de los principales rubros que deben reflejarse en el presupuesto de inversión para el proceso de realizar una investigación:

- *Honorarios de los investigadores*: en el caso de los trabajos de grado se refiere al costo del tiempo que dedican los autores de la investigación, y el valor se obtiene de multiplicar el valor de la hora de cada persona (valor del salario mensual* % número de horas de trabajo por mes) por el número de horas totales que dedicará durante el tiempo que dure la investigación. En el caso de las investigaciones de los investigadores profesionales, el rubro se descompone entre: investigador principal, coinvestigador y auxiliar de investigación.
- *Adquisición de equipos*: cuando para realizar la investigación se requiere la compra de equipos, incluido equipo de computación o equipos especializados para el desarrollo del trabajo de campo.
- *Material de laboratorio*: se requiere cuando es necesario utilizar implementos de laboratorio para realizar el estudio.
- *Material bibliográfico*: son los costos en los que se incurrirá para la obtención de libros, revistas, acceso a bases de datos, etcétera, en el tema de la investigación.
- *Viajes*: se utiliza cuando para realizar el estudio se requiere la movilización de personal investigador o sujetos objeto del estudio a sitios diferentes de la sede de la investigación. Este rubro incluye tiquetes y viáticos cuando sea necesario.
- *Papelería*: son los costos requeridos para la papelería necesaria para realizar el estudio.
- *Imprevistos*: en este rubro se muestran los costos en los que se puede incurrir como consecuencia de contingencias durante el desarrollo de la investigación.

* *En el caso de los estudiantes de pregrado que no suelen tener salario mensual, se toma como referente el salario promedio de un recién egresado en su respectiva profesión.*

www.FreeLibros.me

La Tabla 7.18 ilustra un ejemplo de un presupuesto de inversión para la realización de una investigación.

TABLA 7.18 **Presupuesto de inversión**

Detalle del presupuesto total					
Rubros		Fuentes y usos			Total
		Recursos externos nacionales	Recursos internacionales	Recursos propios	
1	Honorarios de personal	$ 20 000 000		$ 8 000 000	$ 28 000 000
2	Equipos		$ 8 000 000		8 000 000
3	Materiales		1 000 000		1 000 000
4	Salidas de campo			2 000 000	2 000 000
5	Viajes			1 000 000	1 000 000
6	Bibliografía			1 500 000	1 500 000
7	Software			200 000	200 000
8	Publicaciones			1 500 000	1 500 000
9	Honorarios y servicios técnicos			3 000 000	3 000 000
10	Administración			1 000 000	1 000 000
11	Otros			1 000 000	1 000 000
Total		$ 20 000 000	$ 9 000 000	$ 18 200 000	$ 47 200 000

El ejemplo ilustra sobre los rubros en los que se requiere inversión para realizar la investigación, igualmente muestra las fuentes de donde se obtienen los recursos, que para el caso son tres (recursos externos nacionales e internacionales y recursos propios) detallando el valor la fuente y el rubro en el que se utilizará. Finalmente, se muestra el total de la inversión por rubro y para el total del proyecto.

7.14.3 Bibliografía

El documento de un anteproyecto debe terminar con la presentación de la bibliografía, una relación alfabética según el primer apellido de los autores de las fuentes documentales consultadas por los autores del respectivo documento.

La presentación de la bibliografía, como de los demás aspectos del contenido del documento del anteproyecto, se debe elaborar utilizando las normas técnicas de estilo de algunos de los siguientes manuales: *Manual de estilo de publicaciones científicas de Chicago* (*Chicago University*, 2003), *Manual de estilo de publicaciones APA* (*American Psichology Association*, 2005), *Manual de estilo de publicaciones científicas* (*Ohio State University*, 2008), entre otros manuales específicos para tal propósito.

www.FreeLibros.me

7.15 Redacción y entrega del informe

Después de haber analizado y discutido los resultados obtenidos en la investigación, es necesario redactar y entregar el informe final a la institución o área encargada de su evaluación.

Para la redacción del informe es indispensable seguir las normas, los criterios o protocolos establecidos y exigidos por la institución a la cual habrá de entregarse el respectivo informe.

Asimismo, para la entrega deben cumplirse los procedimientos previstos para tal efecto. En investigación es usual que cada país y cada institución de educación superior tengan algunos criterios específicos, tanto para la redacción como para la entrega de los informes finales de los proyectos de investigación.

El anexo A que se encuentra al final del libro, sirve de ejemplo general para la presentación de un documento de anteproyecto y proyecto de investigación a nivel general de trabajo de grado.

Nota: complementario a los elementos antes mencionados y constitutivos del anteproyecto, es necesario elaborar el cronograma de actividades, el presupuesto de inversión requerido para realizar la investigación y finalmente la bibliografía consultada para elaborar el respectivo documento de anteproyecto.

El cronograma debe mostrar las actividades por ejecutar con su respectiva duración en el desarrollo de la investigación, y el presupuesto indicará los diferentes costos en los que se incurrirá para la realización de la respectiva investigación, señalando igualmente los aportantes de esos recursos. En el anexo A se encuentra información al respecto.

www.FreeLibros.me

7.16 Conclusiones

En general, el proceso de investigación científica es entendido como un sistema constituido por diversos aspectos interrelacionados entre sí, pero es un proceso que aunque sus aspectos constitutivos son generales, los mismos se deben considerar de forma particular para cada investigación, dependiendo en gran parte del tema, los objetivos y el tipo de investigación que se va a realizar.

También es importante señalar que este proceso de investigación suele desarrollarse en dos fases, una de anteproyecto o propuesta de investigación y otra de trabajo de campo y redacción del informe final (esta fase se ilustra en el capítulo siguiente).

7.17 Componentes de un documento de anteproyecto

La fase de anteproyecto comprende el planteamiento de los aspectos descritos en este capítulo y los componentes de un documento de un anteproyecto son básicamente los siguientes:

- Preliminares
- Texto o cuerpo del anteproyecto
- Complementarios

7.17.1 Preliminares

En el manuscrito del anteproyecto o propuesta de investigación científica los preliminares son la portada y la tabla de contenido.

7.17.2 Texto o cuerpo del anteproyecto

El texto o cuerpo del anteproyecto está conformado por los siguientes elementos:

- Título.
- Problema de investigación (enunciado y formulación).
- Justificación y delimitación de la investigación.
- Objetivos (general y específicos).
- Marco de referencia (teórico, legal, etcétera).
- Tipo de investigación.
- Hipótesis (si las hay).
- Diseño de la investigación (si la investigación es experimental).

www.FreeLibros.me

- Estrategias metodológicas (población y muestra cuando es necesario, fuentes y técnicas para la obtención de la información, procedimiento o fases de la investigación que se va a realizar).
- Cronograma de actividades y presupuesto de inversión.
- Bibliografía consultada.

7.17.3 Complementarios

Los complementarios en el documento de un anteproyecto son aquellos aspectos del anteproyecto considerados como anexos y que pueden ser documentos de soporte como instrumentos de recolección de información u otros documentos de apoyo.

Nota: el anexo A ilustra un ejemplo de propuesta o anteproyecto de investigación que sirve de referencia para elaborar otras propuestas de su interés. El modelo que aquí se presenta es un ejemplo particular y, por tanto, es necesario, en cada caso, tener en cuenta los criterios de las normas técnicas de estilo de redacción exigidas en cada país o de cada institución a donde se va a presentar el respectivo documento de trabajo de grado.

RESUMEN

Toda investigación en su fase de anteproyecto requiere la elaboración de un cronograma y de un presupuesto de inversión para administrar la dedicación de tiempo y los recursos financieros requeridos para el desarrollo de las diferentes actividades previstas en la investigación que se va a realizar.

El cronograma de actividades consiste en describir y programar las actividades requeridas para desarrollar la investigación, según la secuencia y su respectiva demanda de tiempo, por tal razón, para elaborar el cronograma de actividades, lo primero que se requiere es tener claridad sobre cuáles son las actividades que se deben realizar para llevar a cabo el desarrollo de la investigación prevista según los objetivos de la misma (y la hipótesis que se debe probar cuando sea necesario).

De otra parte, debido a que toda investigación demanda recursos financieros, el presupuesto debe mostrar el valor total y detallado de los costos requeridos por el proyecto para su desarrollo. Igualmente, se deben detallar la fuentes (es decir, de dónde se obtendrá el dinero) para su financiación.

Los siguientes son algunos de los principales rubros que se deben reflejar en el presupuesto de inversión para un proyecto de investigación:

www.FreeLibros.me

- Honorarios de los investigadores
- Adquisición de equipos
- Material de laboratorio
- Viajes
- Papelería
- Imprevistos

EJERCICIOS DE REPASO Y ANÁLISIS

1. ¿Qué es un cronograma de actividades en un anteproyecto de investigación científica?
2. ¿Por qué son importantes los presupuestos de inversión en un proyecto de investigación?
3. Elabore un esquema de un presupuesto de inversión para un anteproyecto de investigación, en el que se muestren los principales rubros que demandan recursos financieros.
4. ¿Cuáles son los principales criterios para tener en cuenta en la elaboración de la bibliografía de un documento de anteproyecto de investigación?
5. Enuncie o mencione los principales componentes de un documento de anteproyecto de investigación.

www.FreeLibros.me